Characterization of human ALKBH4 – an AlkB homolog with a possible function in gene regulation

Thesis for the degree of *Philosophiae Doctor*

by

Linn Grimsdatter Bjørnstad

NORWEGIAN **CANCER** SOCIETY

Department of Biosciences,
Faculty of Mathematics and Natural Sciences,
University of Oslo, Norway

2013

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 1361

ISSN 1501-7710

Cover: Inger Sandved Anfinsen.
Printed in Norway: AIT Oslo AS.

Produced in co-operation with Akademika Publishing.
The thesis is produced by Akademika publishing merely in connection with the thesis defence. Kindly direct all inquiries regarding the thesis to the copyright holder or the unit which grants the doctorate.

Table of Contents

Acknowledgements

The work presented in this thesis was performed at the Department of Molecular Biosciences (now Department of Biosciences), University of Oslo, from 2008 to 2012. Financial support for the study was provided by the Norwegian Cancer Society.

First of all, I would like to thank my supervisor, Professor Pål Falnes, for introducing me to an interesting field of research and for giving me the opportunity to do my Ph.D. in his research group. Thank you for always being available to discuss the project and for your valuable comments in the writing process.

Furthermore, I am grateful to my co-supervisor Trine Johansen Meza for practical assistance in the lab as well as encouraging discussions. I also thank Leonardo Meza-Zepeda for stepping in as my co-supervisor when Trine left the lab, an offer that was greatly appreciated. I also thank both of you for critical reading of this thesis.

I would also like to thank Giorgio Zoppellaro and Kristoffer Andersson for introducing me to EPR spectroscopy. Kristoffer, I also thank you for critical reading of this thesis.

I would like to thank all co-authors for their invaluable contributions to the two publications that have emerged from this work.

Thanks a lot to all former and present group members of the Falnes lab for creating a casual and positive research environment. Fellow Ph.D. students and other colleagues at the department also deserve thanks for our conversations and coffee breaks as well as for their efforts in trying to keep me motivated during the writing process.

I also wish to thank my family and friends for always supporting me. Especially, thank you, Ole, for all your love and understanding. I would never have finished this without you backing me. Last, but not least, thank you, Tage, for being the best little man, and for reminding me of what is really important in life.

Blommenholm, February 2013
Linn Grimsdatter Bjørnstad

Abbreviations

1-meA	1-methyladenine
1-meG	1-methylguanine
2OG	2-oxoglutarate
3-meC	3-methylcytosine
3-meT	3-methylthymidine
3-meU	3-methyluridine
5-caC	5-carboxylcytosine
5-fC	5-formylcytosine
5-hmC	5-hydroxymethylcytosine
5-hmU	5-hydroxymethyluracil
5-meC	5-methylcytosine
6-meA	N^6-methyladenine
AF9	ALL1-fused gene from chromosome 9 protein
AID/APOBEC	activation induced deaminase/apolipiprotein B mRNA editing cytokine deaminase
Ala (A)	alanine
ALKBH	AlkB homolog
AP	apurinic/apyrimidinic (abasic)
Arg (R)	arginine
Asn (N)	asparagine
Asp (D)	aspartic acid
ATBF1	AT-motif binding factor 1
ATP	adenosine triphosphate
Base J	β-D-glucopyranosyloxy-methyluracil
BER	base excision repair
BLAST	basic local alignment search tool
BRD	bromodomain
CBP	cAMP response element-binding protein (CREB)-binding protein
CGI	CpG island
cm^5U	5-carboxymethyluridine
CpG	cytosine-phosphate-guanine
DBD	DNA binding domain
DNA	deoxyribonucleic acid
DNMT	DNA methyltransferase
DSB	DNA double-strand break
DSBH	double-stranded beta-helix
dsDNA	double-stranded DNA
E. coli	*Escherichia coli*
ENL	eleven-nineteen leukemia
EPR	electron paramagnetic resonance
Fe(II)	ferrous iron
Fe(III)	ferric iron
Fe(IV)	ferryl iron
Fe-S	iron-sulfur
FGF	fibroblast growth factor
FIH	factor inhibiting HIF
FTO	fat mass and obesity-associated protein
GEO	Gene Expression Omnibus
Glu (E)	glutamic acid
Gly (G)	glycine
GO	gene ontology
H3K4	lysine 4 in histone 3
H3K79	lysine 79 in histone H3
HAT	histone acetyltransferase
HDAC	histone deacetylase
HDM	histone demethylase
HIF	hypoxia-inducible factor
His (H)	histidine
HMT	histone methyltransferase
HRR	homologous recombination repair
HSF4	heat-shock transcription factor 4
JBP	J-binding protein
JmjC	jumonji C
Leu (L)	leucine
Lys (K)	lysine
MBP	methyl-CpG-binding protein
mcm^5s^2U	5-methoxycarbonylmethyl-2-thiouridine
mcm^5U	5-methoxycarbonylmethyl-uridine
mcm^5Um	5-methoxycarbonylmethyl-2'-*O*-methyluridine
MEF	mouse embryonic fibroblast
MLL	mixed-lineage leukemia
MMS	methyl methanesulphonate
mRNA	messenger RNA
MT	methyltransferase
NCBI	National Center for Biotechnology Information
NO	nitric oxide
PADI	peptidyl arginine deiminase
PAFc	polymerase-associated factor complex
PHD	plant homeodomain
PHD	prolyl hydroxylase domain
Phe (F)	phenylalanine
Pol II	RNA polymerase II

PRMT	protein arginine methyltransferase
Pro (P)	proline
RA	retinoic acid
RNA	ribonucleic acid
RRM	RNA recognition motif
rRNA	ribosomal RNA
S. cerevisiae	*Saccharomyces cerevisiae*
SAM	S-adenosyl methionine
Sec (U)	selenocysteine
SEC	super elongation complex
Ser (S)	serine
SET	Su(var)3-9-Enhancer of zeste-Trithorax
siRNA	small interfering RNA
(*S*)-mchm^{5}U	5-[methoxycarbonylhydroxy-methyl]-uridine
ssDNA	single-stranded DNA
ssRNA	single-stranded RNA
TauD	Taurine dioxygenase D
TDG	thymine DNA glycosylase
TES	testin
TET	ten-eleven translocation
TfdA	2,4-dichlorophenoxyacetate dioxygenase A
tRNA	transfer RNA
Trp (W)	tryptophan
Tyr (Y)	tyrosine
UV-vis	ultraviolet-visible
Y2H	yeast two-hybrid
YEATS	Yaf9-ENL-AF9-Taf14-Sas5

List of papers

Paper I

Bjørnstad L. G., Zoppellaro G., Tomter A. B., Falnes P. Ø. and Andersson K. K.

Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4

Biochem. J. (2011) **434**, 391-398

Paper II

Bjørnstad L. G., Meza T. J., Otterlei M., Olafsrud S. M., Meza-Zepeda L. A. and Falnes P. Ø.

Human ALKBH4 interacts with proteins associated with transcription

PLoS ONE (2012) **7**(11): e490945. doi:10.1371/journal.pone.0049045

Introduction

Iron (Fe(II))- and 2-oxoglutarate (2OG)-dependent dioxygenases

The iron and 2-oxoglutarate (Fe(II)/2OG)-dependent dioxygenases constitute a large family of enzymes which catalyze oxidation reactions. Characteristic for this family is their catalytic requirement for non-heme ferrous iron (Fe(II)) as co-factor and 2-oxoglutarate (2OG) as co-substrate. Crystallographic structure analyses of multiple family members have revealed these enzymes to possess a common core fold, the double-stranded beta-helix (DSBH) fold (Figure 1A), which forms a distorted, barrel-shaped structure from two antiparallel β-sheets (reviewed in McDonough et al. (2010)). In most cases, additional secondary structure elements or domains are involved in substrate binding. Such elements include subfamily-specific inserts between the βIV and βV strands or, less frequently, other β-strands. The length of these inserts vary from a few amino acids (Elkins et al., 2003), to fully functional domains (Elkins et al., 2002). Similarly, additional elements or domains are also commonly found N- or C-terminally to the core. In the more open end of the barrel, the DSBH core comprises a conserved Fe(II)-chelating 2-His-1-carboxylate active site motif, in which the carboxylate is from the side chain of aspartic acid or, less frequent, glutamic acid (Hegg and Que 1997). This motif is known as the "facial triad" because the three iron coordination sites are arranged to constitute one octahedral face (Hegg and Que 1997). The three remaining sites are left to accommodate other endogenous ligands or exogenous ligands originating from water, 2OG, prime substrate or O_2. The large flexibility of the Fe(II)/2OG dioxygenase fold, reflected by the structure of the facial triad, explains the great versatility of this enzyme family with respect to chemical transformations and substrate recognition.

The consensus catalytic mechanism of Fe(II)/2OG-dependent dioxygenases (Figure 1B) involves initial binding of 2OG to the active site Fe(II) in a bidentate fashion, leading to displacement of two water molecules. This is thought to facilitate binding of prime substrate and, upon subsequent binding of O_2, oxidative decarboxylation of 2OG results in a reactive Fe(IV)-oxo species which accomplishes substrate oxidation with reduction of Fe(IV) to Fe(II). Dissociation of the oxidized product, the co-product succinate and CO_2 completes the catalytic cycle. However, the detailed reaction mechanisms can vary slightly between different family members, even between two enzymes acting on the same substrate (Neidig et

al., 2006). In animals, Fe(II)/2OG-dependent dioxygenases perform *C*-hydroxylations or hydroxylation-mediated *N*-demethylations of amino acid side chains and nucleic acid bases, as well as lipid hydroxylations. In microorganisms and plants, they have also been reported to perform more complex oxidation reactions such as cyclizations (Zhang et al., 2000), desaturations (Clifton et al., 2003; Sleeman et al., 2004) and halogenations (Vaillancourt et al., 2005; Vaillancourt et al., 2005) in addition to the more common hydroxylation reactions.

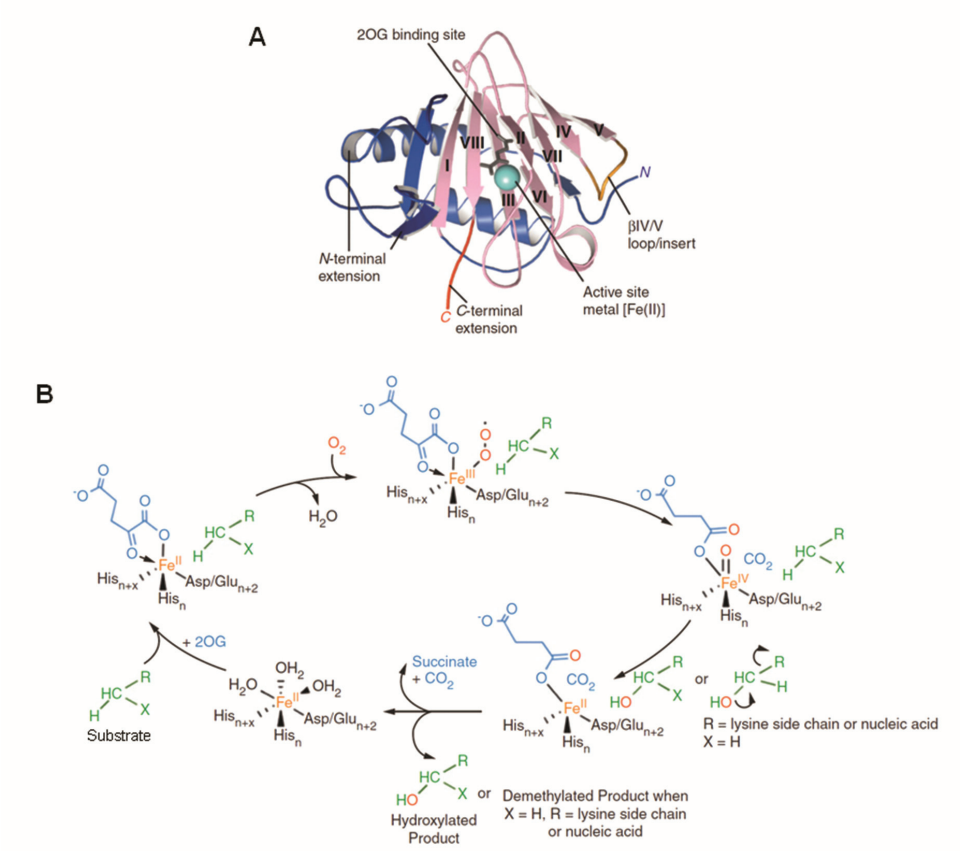


Figure 1. Structure and catalytic mechanism of Fe(II)/2OG-dependent dioxygenases. A) General structure of the double-stranded beta-helix (DSBH) core fold. **B)** Binding of primary substrate is preceded by 2-oxoglutarate (2OG) binding. Subsequent binding of O_2 and oxidative decarboxylation of 2OG results in a reactive Fe(IV)=O intermediate which oxidizes the substrate, while reducing Fe(IV) to Fe(II). Release of succinate and hydroxylated product enables successive reaction cycles. Adapted from Aik et al. (2012).

As a result of the diversity of the chemical reactions catalyzed by Fe(II)/2OG-dependent dioxygenases, such enzyme activity is central in a great number of biological processes. In addition to their well-known roles in fatty acid metabolism (Hulse et al., 1978) and post-translational modification of collagen (Hutton et al., 1966; Hutton et al., 1967), recent findings have also linked these enzymes to hypoxic signaling (Ivan et al., 2001; Jaakkola et al., 2001), DNA repair (Falnes et al., 2002; Trewick et al., 2002), epigenetic gene regulation (Tsukada et al., 2006; Yu et al., 2007; Tahiliani et al., 2009), mRNA splicing (Jia et al., 2011; Zheng et al., 2012) and tRNA hypermodification (Fu et al., 2010; van den Born et al., 2011).

Here, the functions of some of the proteins in this fascinating enzyme family will be described in further detail.

AlkB proteins

E. coli AlkB

The *alkB* gene was first identified in a screen isolating *Escherichia coli (E. coli)* mutants sensitive to the S_N2 alkylating agent methyl methanesulphonate (MMS) (Kataoka et al., 1983). Although initial progress was made in terms of cloning the gene and purifying the encoded protein (Kataoka and Sekiguchi 1985), the precise role of AlkB in alkylation resistance remained enigmatic. However, the original assumption that AlkB *per se* could be an alkylation damage repair enzyme was supported when human cell lines over-expressing AlkB was shown to display increased resistance to MMS (Chen et al., 1994). A leap in direction of elucidating the AlkB function came with the finding that a*lkB* mutants were defective in reactivation of MMS-inactivated single-stranded, but not double-stranded, DNA bacteriophages (Dinglay et al., 2000). As the N1 position in purines and N3 position in pyrimidines are engaged in Watson-Crick base-pairing, and thus protected from the effect of S_N2 agents in double-stranded DNA, this indicated that the base lesions 1-methyladenine (1-meA) and 3-methylcytosine (3-meC) could be the AlkB substrates. Soon after, the ultimate clue to identification of the AlkB function was provided when a bioinformatics report placed the protein in the superfamily of Fe(II)- and 2-oxoglutarate (2OG)-dependent dioxygenases (Aravind and Koonin 2001). Consequently, nearly two decades after the initial identification of AlkB, two groups independently determined its function in repair of 1-meA and 3-meC lesions in single-stranded DNA (Falnes et al., 2002; Trewick et al., 2002).

AlkB catalyzes oxidation of the aberrant methyl group, resulting in an unstable hydroxymethyl group which is spontaneously released as formaldehyde, thereby regenerating the undamaged base (Figure 2). This oxidative demethylation reaction requires ferrous iron as co-factor and 2-oxoglutarate as co-substrate, and the oxidizing agent molecular oxygen. Decarboxylation of 2OG to succinate and CO_2 is coupled to oxidation of the primary substrate. The discovery of the AlkB function revealed a novel mechanism for direct reversal of DNA alkylation damage.

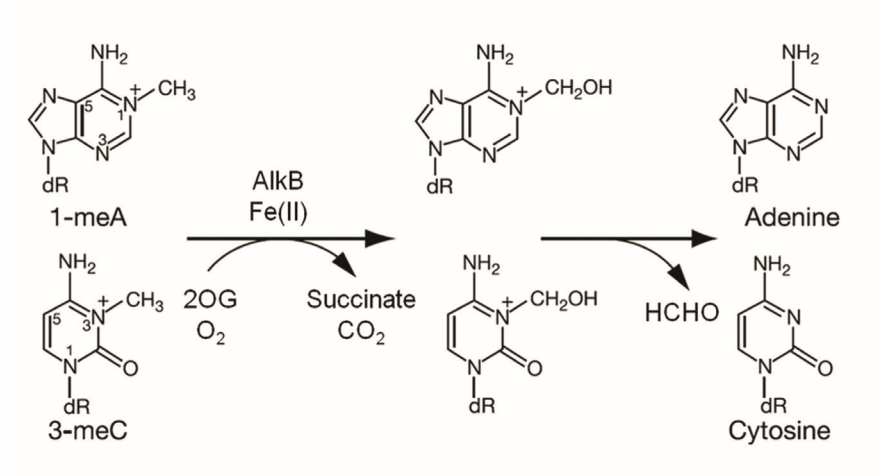


Figure 2. Reaction mechanism of *E. coli* AlkB. AlkB reverses damaging alkyl groups in the N1-position of purines and the N3-position of pyrimidines, here illustrated by 1-meA and 3-meC, in DNA or RNA. The methyl group is oxidized in presence of Fe(II), 2-oxoglutarate (2OG) and O_2, resulting in its destabilization and spontaneous release as formaldehyde (HCHO), thereby regenerating the undamaged base. Decarboxylation of 2OG to succinate and CO_2 is coupled to substrate oxidation. Adapted from Trewick et al. (2002).

The crystal structure of AlkB in complex with a single-stranded, 1-meA-containing DNA trinucleotide, Fe(II) and 2-oxoglutarate confirmed the structural prediction by Aravind and Koonin that its catalytic core adapts the DSBH fold typical for the Fe(II)/2OG-dependent dioxygenases (Aravind and Koonin 2001; Yu et al., 2006). This structure further verified the presence of two motifs predicted by Aravind and Koonin: the Fe(II)-chelating $HXDX_nH$ triad (His131, Asp133 and His187), as well as the AlkB-specific RX_5R motif (Arg204 and Arg210) which is involved in coordination of 2OG (Aravind and Koonin 2001; Yu et al., 2006).

Subsequently, another AlkB structure, complexed with double-stranded DNA, gave further insight into the means by which AlkB binds its substrate (Yang et al., 2008). This report revealed the basis for the preference of AlkB for single-stranded substrates, as the enzyme primarily interacts with the damaged strand. Moreover, AlkB-mediated compression of the two bases flanking the lesion site leads to their stacking, as well as facilitating flipping of the base to be repaired into the active enzyme site.

Although less efficiently, AlkB also acts on the structural 1-meA and 3-meC analogs 1-methylguanine (1-meG) and 3-methylthymine (3-meT), thus suggesting the former to be the primary substrates (Delaney and Essigmann 2004; Falnes 2004; Koivisto et al., 2004). Subsequent reports have further extended the substrate repertoire to also comprise bulkier lesions, as AlkB also acts on ethyl- and propyl-lesions, as well as exocyclic etheno- and ethano-adducts (Koivisto et al., 2003; Delaney et al., 2005; Frick et al., 2007). In addition to this rather broad substrate specificity in DNA, AlkB is also active on RNA (Aas et al., 2003). AlkB-mediated RNA demethylation was further shown to reactivate MMS-inactivated RNA bacteriophages, thus demonstrating that AlkB also displays *in vivo* repair activity towards alkylated RNA (Aas et al., 2003). Another report, further supporting biological importance of RNA repair, showed that AlkB can functionally recover both mRNA and tRNA molecules previously inactivated by chemical methylation (Ougland et al., 2004). These intriguing discoveries have inevitably revived the debate concerning the biological role of RNA repair, a debate which is still ongoing.

The *alkB* gene is part of the so-called adaptive response to alkylation damage, which also includes three other genes, *ada*, *alkA* and *aidB* (reviewed in Sedgwick and Lindahl (2002)). Like AlkB, both Ada and AlkA are DNA repair proteins (McCarthy et al., 1984; Demple et al., 1985). AidB has been suggested to be involved in detoxification of alkylating agents rather than in repair of the damaged DNA (Landini et al., 1994). Positive regulation of the adaptive response is provided by Ada, an alkyltransferase which mediates non-enzymatic transfer of damaging methyl groups from DNA to certain cysteine residues within Ada itself (Lindahl et al., 1982). Methylated Ada then acts as a transcription factor to induce expression of the regulon comprising all four adaptive response genes (reviewed in Landini and Volkert (2000)).

Mammalian AlkB homologs (ALKBHs)

Based on sequence similarity, nine homologs of the *E. coli* AlkB protein, ALKBH1-8 and FTO, have so far been identified in mammals (Kurowski et al., 2003; Gerken et al.,

2007). These proteins all display a double-stranded beta-helix (DSBH) jelly roll fold similar to that of AlkB and Fe(II)/2OG-dependent dioxygenases in general. However, the sequence homology is mainly restricted to the AlkB specific catalytic core motif, $HXDX_nH…RX_5R$, which is conserved in the mammalian homologs. The phylogenetic relationship of ALKBH proteins is shown in Figure 3. A brief description of each homolog is given below, and their main substrates and functions are summarized in Table 1.

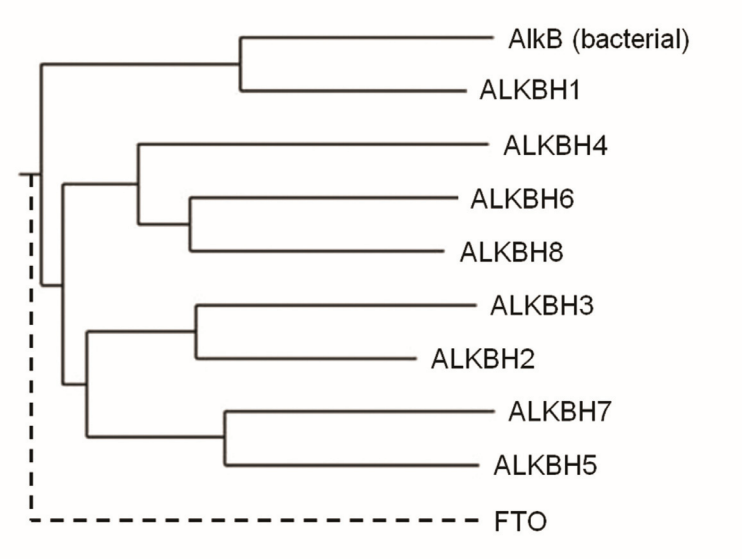


Figure 3. Phylogenetic relationship between mammalian ALKBH proteins. The more distantly related FTO protein is indicated with a dotted line. Adapted from Kurowski et al. (2003).

ALKBH1

ALKBH1, the homolog with the highest similarity to *E. coli* AlkB was also the first mammalian AlkB protein to be discovered (Wei et al., 1996). Despite the high sequence similarity between the two, a function of ALKBH1 in DNA repair is debated. The initial report suggested it to be a functional AlkB homolog since its expression partially rescued *alkB* mutant *E. coli* cells from MMS-induced cell death (Wei et al., 1996). Supporting this, another group reported this enzyme to be a functional mitochondrial AlkB homolog possessing 3-meC demethylase activity towards both DNA and RNA (Westbye et al., 2008). However, this DNA repair activity was very weak and never reproduced by others. Interestingly, two studies using gene-targeted mice rather pointed towards a role of ALKBH1 in embryogenic gene regulation (Pan et al., 2008; Nordstrand et al., 2010). In the first of

these, disruption of placental trophoblast differentiation was observed in *Alkbh1*-targeted mice (Pan et al., 2008). Additionally, they showed localization of Alkbh1 to nuclear euchromatin as well as its binding to the placental gene repressor Mrj which exerts its effect through recruitment of type II histone deacetylases (HDACs). Mutually exclusive interaction of Alkbh1 and HDAC5 to Mrj thus suggested a stimulatory effect of Alkbh1 on placental trophoblast lineage differentiation. The second study revealed *Alkbh1*-deficient mice to display sex-ratio distortion against females, unilateral developmental defects in the eye and skeleton, and, consistently, embryonic misexpression of bone morphogenetic proteins (Bmps), thus Alkbh1 was suggested to be a histone demethylase functioning during embryogenesis and spermatogenesis (Nordstrand et al., 2010). Very recently, this controversial proposal of ALKBH1 possessing a histone substrate was supported, as histone H2A was demonstrated to stimulate decarboxylation of 2-oxoglutarate by recombinant Alkbh1, and H2A purified from *Alkbh1*-deficient, but not wild-type, mouse embryonic fibroblast (MEF) cells displayed dimethyl groups on certain lysine residues (Ougland et al., 2012). Consistently, repression of a Piwi-interacting RNA (piRNA) cluster in spermatocytes by Alkbh1 and the testis-specific transcription repressor Tzfp was also reported, in which the contribution of Alkbh1 was suggested to be chromatin structure alteration by demethylation of histone H2A (Nordstrand et al., 2012).

In addition to its demethylase activity, ALKBH1 possesses *in vitro* DNA lyase activity, cleaving DNA at abasic (AP) sites in both ssDNA and dsDNA to generate strand breaks (Muller et al., 2010). This activity is not dependent on Fe(II) and 2OG, suggesting it to be ascribed to a second active site. The biological relevance of this activity is, however, not known.

ALKBH2 and ALKBH3

With the identification of ALKBH2 and ALKBH3, humans were shown to possess two proteins with the ability to complement reactivation of an MMS-inactivated single-stranded DNA bacteriophage in an *alkB* mutant *E. coli* strain as well as to remove 1-meA and 3-meC from single-stranded DNA by oxidative demethylation (Duncan et al., 2002; Aas et al., 2003). Different substrate specificities have been shown for the two homologs, and while ALKBH2 prefers double-stranded substrates, single-stranded ones are favored by ALKBH3 (Aas et al., 2003; Falnes et al., 2004). Furthermore, while ALKBH2 is exclusively active on DNA, ALKBH3 also reverses alkylation damage in RNA substrates (Aas et al., 2003). In addition, ALKBH3 was, like *E. coli* AlkB, able to both reactivate the MMS-alkylated single-

stranded RNA phage MS2 in *alkB* deficient *E. coli* cells and recover the functions of alkylation-inactivated mRNA and tRNA molecules in the processes of translation and aminoacylation (Ougland et al., 2004). Notably, AlkB proteins have also been identified in viruses with RNA genomes, further supporting a biological significance of RNA repair (Aravind and Koonin 2001).

The crystal structures of ALKBH2 and AlkB in complex with dsDNA revealed that ALKBH2, in contrast to AlkB, interacts extensively with the non-damaged strand, thus explaining the basis of the ALKBH2 preference for double-stranded substrates (Yang et al., 2008). Additionally, structural differences were found in the base flipping mechanisms of the two enzymes: while AlkB distorts the DNA backbone to allow stacking of the residues flanking the damage site, the corresponding base gap is filled by intercalation of an aromatic residue (Phe102) in ALKBH2 (Yang et al., 2008). Furthermore, the structure of ALKBH3 in absence of DNA/RNA has been solved (Sundheim et al., 2006). However, its superimposition onto those of AlkB and ALKBH2 surprisingly revealed the structure of the ALKBH3 active site to be most similar to that of ALKBH2, as both human proteins, but not AlkB, display a divergent β-hairpin loop close to the substrate binding groove (Sundheim et al., 2008; Yang et al., 2008). The combination of motif-swapping and mutagenesis analysis demonstrated that the hairpin is a determinant of substrate specificity, and acidic residues (Arg101-Glu102-Asp103) in the ALKBH3 hairpin provide the structural basis for its single-strand preference, while the responsibility for stabilization and base-flipping of double-stranded ALKBH2 substrate was ascribed to a hydrophobic network partially composed of the hairpin loop (Chen et al., 2010; Monsen et al., 2010). Notably, the absence of this hairpin in AlkB might explain its broader substrate specificity relative to ALKBH3 (Mishina et al., 2005).

Moreover, the ALKBH3 structure further revealed auto-hydroxylation of an essential leucine residue (Leu177) in the active site, which was proposed to serve as an inactivating switch to prevent formation of damaging oxygen radicals in the absence of substrate (Sundheim et al., 2006).

Distinct biological roles of ALKBH2 and ALKBH3 are indicated not only by their substrate preferences, but also by their different subcellular localization patterns. While ALKBH2 is generally distributed evenly throughout the nucleoplasm, cell-cycle dependent relocation to replication foci occurs in S-phase cells (Aas et al., 2003). ALKBH3, however, localizes mainly to the nucleoplasm, with partial cytoplasmic occurrence as well as general exclusion from nucleoli. Thus, the two proteins likely function in DNA repair during replication and transcriptional maintenance of single-stranded DNA/RNA, respectively, and

appear to together cover the cellular functions performed by AlkB in *E. coli* (Aas et al., 2003). Supporting this suggestion of ALKBH2 being the main enzyme to reverse genomic 1-meA and 3-meC, increased accumulation of 1-meA under normal physiological conditions was observed in *Alkbh2*-targeted mice with respect to wild-type as well as *Alkbh3*-targeted mice (Ringvoll et al., 2006). Furthermore, *Alkbh2*-, but not *Alkbh3*-targeted, MEF cells displayed reduced repair kinetics after MMS-treatment (Ringvoll et al., 2006), and only ALKBH2 complemented reversal of etheno adducts in *alkB* deficient *E. coli* cells (Ringvoll et al., 2008). However, in addition to the role of ALKBH3 in RNA repair, the enzyme was recently suggested to function in biological maintenance of genome integrity in a subset of cancer cell lines in which it reverses 3-meC in dsDNA previously unwinded, and thus made single-stranded, by the 3'-5' DNA helicase activating signal cointegrator complex subunit 3 (ASCC3) to ensure proliferation (Dango et al., 2011).

ALKBH4-ALKBH7

Four homologs, ALKBH4, ALKBH5, ALKBH6 and ALKBH7, have all been scarcely studied and their functions are still unknown. These proteins are, based on primary sequence, more similar to ALKBH2/ALKBH3 than to *E. coli* AlkB/ALKBH1 (Kurowski et al., 2003) (Figure 3). Moreover, ALKBH5/ALKBH7 and ALKBH4/ALKBH6/ALKBH8 are in-paralogs, meaning they originated from gene duplication subsequent to the initial radiation of the AlkB lineage (Kurowski et al., 2003). In an initial functional study of these homologs, no *in vitro* activity towards 1-meA and 3-meC was detected for the successfully purified recombinant human ALKBH4, ALKBH6 and ALKBH7 proteins (Lee et al., 2005). However, subcellular localization to the nucleus as well as the cytoplasm has been reported for green fluorescent protein (GFP) fusions of all four homologs, indicating the possibility of nuclear functions (Tsujikawa et al., 2007). Furthermore, ALKBH5 has previously, together with ALKBH6, been suggested to target nucleic acids, as they, like the AlkB proteins known to repair DNA, display high pI values which are compatible with and potentially important for DNA association (Sedgwick et al., 2007). Likewise, the relatively low pI values of ALKBH4 and ALKBH7 could possibly disfavour their DNA binding abilities, and these proteins were consequently proposed to rather act on proteins (Sedgwick et al., 2007). We recently demonstrated uncoupled 2-oxoglutarate decarboxylase activity of human ALKBH4 in absence of prime substrate, suggesting this protein to be a *bona fide* Fe(II)/2OG-dependent dioxygenase (Bjornstad et al., 2011). Another recent study reported uncoupled 2-oxoglutarate decarboxylation by ALKBH5, which was further indicated to possess a regulatory role in the

cellular hypoxia response, as *ALKBH5* was the only *ALKBH* gene to be induced by hypoxia inducible factor-1α (HIF-1α) under hypoxic conditions (Thalhammer et al., 2011). Very recently, supporting the previously proposed activity of ALKBH5 towards nucleic acids, the protein was shown to demethylate N^6-methyladenine (6-meA) in mRNA both *in vitro* and *in vivo* (Zheng et al., 2012). *ALKBH5* deficiency was further demonstrated to perturb mRNA export, RNA metabolism and mRNA processing factor assembly in addition to spermatogenesis (Zheng et al., 2012). In conclusion, it appears likely that at least some of the uncharacterized homologs are involved in processes other than DNA repair, and that the target repertoire of the ALKBH protein family will expand in the future to include proteins in addition to nucleic acids.

ALKBH8

As the only AlkB homolog, ALKBH8 comprises two annotated functional domains in addition to the AlkB domain, these being an RNA recognition motif (RRM) and an S-adenosyl methionine (SAM)-dependent methyltransferase (MT) domain (Tsujikawa et al., 2007). Moreover, ALKBH8 is also the only one displaying an exclusively cytoplasmic localization (Tsujikawa et al., 2007). The MT domain displays sequence similarity to the *S. cerevisiae* Trm9 protein, which has previously been shown to catalyze methylesterification of modified uridine nucleotides in the wobble position of the anti-codon loops of $tRNA^{Gly}$ and $tRNA^{Arg}$, thus generating the 5-methoxycarbonylmethyluridine (mcm^5U) modification as well as its derivative 5-methoxycarbonylmethyl-2-thiouridine (mcm^5s^2U) (Kalhor and Clarke 2003). *Alkbh8*-targeted mice lack both these tRNA modifications as well as the 2'-*O*-ribose-methylated form of mcm^5U (mcm^5Um) (Songe-Moller et al., 2010). Through investigation of recombinant ALKBH8 domains, the ALKBH8-MT activity was shown to be responsible for the final methylation step in generation of the mcm^5U modification, and this activity was further shown to strictly depend on TRM112, a small protein subunit also required for the function of other tRNA and protein methyltransferases (Fu et al., 2010; Songe-Moller et al., 2010).

In mammals, the mcm^5U and mcm^5Um modifications exist also in $tRNA^{Sec}$, which decodes the UGA stop codon as the so-called 21st amino acid selenocysteine (Sec) during translation of selenoproteins (reviewed in Hatfield et al. (2006)). The mcm^5Um/mcm^5U ratio is likely important for efficient selenoprotein synthesis, and consistently, the *Alkbh8*-targeted mice showed decreased expression of the selenoprotein GPx1 (Songe-Moller et al., 2010).

Soon after the elucidation of the ALKBH8-MT activity, a recombinant Alkbh8 version in which the MT domain was deleted (RRM-AlkB) was demonstrated to hydroxylate mcm^5U into (*S*)-$mchm^5U$ in a synthetic $tRNA^{Gly}$ anticodon loop-resembling substrate (Fu et al., 2010). Furthermore, using mice engineered to separately express the two domains (MT or AlkB), the AlkB domain was found to hydroxylate mcm^5U previously generated from cm^5U by the MT domain, resulting in the (*S*)-$mchm^5U$ modification in $tRNA^{Gly}$ (van den Born et al., 2011). Thus, the MT and AlkB domains of ALKBH8 act sequentially to first generate the wobble position modification mcm^5U in a subset of tRNA isoacceptors, which in $tRNA^{Gly}$ is subsequently hydroxylated to (*S*)-$mchm^5U$.

Recently, increased insight into the RNA binding properties of ALKBH8 was provided with publication of the crystal structure of its RRM-AlkB moiety, demonstrating structural stabilization of the AlkB domain by a Zn(II)-binding cysteine cluster in its C-terminus as well as formation of a unified interaction surface by the RRM and AlkB domains, mediating contact with the tRNA substrate (Pastore et al., 2012). Through additional thermodynamic and crystallographic analysis, the authors found the RRM motif together with an N-terminal α-helix preceding the RRM to contribute to unspecific RNA binding, while substrate specificity is likely rather provided by a structural fold spanning both the RRM motif and the AlkB domain (Pastore et al., 2012).

In conclusion, the elucidation of ALKBH8 as a multifunctional RNA modification enzyme was the first definite evidence of ALKBH function extending beyond DNA repair.

FTO

A bioinformatics analysis of the fat mass- and obesity-associated (FTO) protein recently resulted in expansion of the mammalian AlkB family to include this protein as the ninth member. FTO was predicted to possess the characteristic DSBH fold containing both the iron-binding $HXDX_nH$ triad and the 2-oxoglutarate-coordinating RX_5R motif in addition to a C-terminal domain without homology to any known protein sequence (Gerken et al., 2007). FTO is associated with predisposition to obesity, which is explained by the presence of a common polymorphism in the first intron of the *FTO* gene (Frayling et al., 2007). The molecular mechanism behind the association of FTO with susceptibility to obesity has not yet been clarified, although a function in regulation of energy homeostasis has been proposed (Fischer et al., 2009). Such a function was suggested on the basis of the observation that *Fto* deficient mice displayed a lean phenotype with reduced fat mass, despite a relative increase in food intake and unchanged locomotor activity in comparison to wild-type mice (Fischer et al.,

2009). Similarly, mice with higher *Fto* copy-number are obese, however, primarily as a result of increased food intake, with no significant changes in energy expenditure and locomotor activity (Church et al., 2010). Consistent with a role in nucleic acid demethylation, similar to the two functional AlkB homologs ALKBH2 and ALKBH3, FTO localizes to the nucleus (Gerken et al., 2007). Initial studies on its enzymatic activity reported FTO to display weak *in vitro* demethylase activity towards single-stranded, but not double-stranded, oligonucleotides containing 3-meT or 3-meU (Gerken et al., 2007; Jia et al., 2008). A slightly higher activity towards the methylated uridine- compared to thymidine-containing substrate, suggested that the biological substrate could be RNA (Jia et al., 2008).

Explaining the preference for single-stranded substrates, the crystal structure of FTO in complex with 3-meT showed the presence of an additional substrate recognition loop, not present in AlkB, which would be in sterical conflict with the non-methylated strand (Han et al., 2010). The crystal structure also revealed the critical function of the FTO C-terminal extension in stabilization of the N-terminal domain to ensure its catalytic activity.

Recently, Jia and co-workers described efficient *in vitro* oxidative demethylase activity by FTO towards N^6-methyladenosine (6-meA) in synthetic ssDNA/RNA substrates (Jia et al., 2011). RNA was further experimentally demonstrated to be a biologically relevant target, as the 6-meA levels in cellular mRNA decreased in response to FTO over-expression, and a corresponding decrease was observed upon siRNA-mediated knock-down of *FTO* (Jia et al., 2011). Although the 6-meA modification is ubiquitously found in nuclear mRNAs, its

Table 1. Overview of main substrates and functions of mammalian ALKBH proteins.

Enzyme	Main substrate(s)	Function	References
ALKBH1	3-meC in ssDNA/RNA Methyl-lysine in H2A	DNA/RNA repair Gene regulation	Westbye et al., 2008 Ougland et al., 2012
ALKBH2	1-meA and 3-meC in dsDNA	DNA repair	Aas et al., 2003 Falnes et al., 2004
ALKBH3	1-meA and 3-meC in ssDNA/RNA	DNA/RNA repair	Aas et al., 2003 Falnes et al., 2004
ALKBH4	?	?	
ALKBH5	6-meA in mRNA	mRNA export	Zheng et al., 2012
ALKBH6	?	?	
ALKBH7	?	?	
ALKBH8	cm^5U in tRNA mcm^5U in tRNA	tRNA modification	Songe-Moller et al., 2010 Fu et al., 2010a van den Born et al., 2011 Fu et al., 2010b
FTO	6-meA in mRNA	pre-mRNA splicing	Jia et al., 2011

biological function is not fully understood. However, a recent report on the 6-meA patterns at the global transcriptome level showed this modification to be fundamental in the process of pre-mRNA splicing (Dominissini et al., 2012). Thus, novel data strongly suggest a functional role of FTO in regulation of gene expression rather than repair.

Gene regulation from the perspective of Fe(II)/2OG-dependent dioxygenases

Several Fe(II)/2OG-dependent dioxygenases display well established functions in gene regulation. Among these are proteins which act through modification of the chromatin structure to regulate transcription. Moreover, increasing evidence implies that other Fe(II)/2OG-dependent dioxygenases probably function in regulation of subsequent steps of gene expression (Figure 4).

Chromatin structure and function

DNA is present in the cell nucleus in the form of chromatin, a nucleoprotein complex which consists of histone proteins in addition to the DNA. The repetitive unit of chromatin is the nucleosome, which is composed of an octamer of two copies of each of the four core histones H2A, H2B, H3 and H4 around which 147 base pairs of DNA are wrapped twice (Luger et al., 1997), and the short stretch of DNA that couples adjacent core particles, the linker DNA, with its bound fifth histone protein, the linker histone H1 (Happel and Doenecke 2009). The positively charged histone proteins form electrostatic interactions with the negative DNA molecule to stabilize the nucleosome from which the unstructured N-terminal tails of the otherwise globular core histones protrude (Luger and Richmond 1998). Furthermore, the linker histone is thought to be involved in stabilization of higher order chromatin structures (Robinson and Rhodes 2006).

The hierarchical structure of chromatin reflects its function in DNA organization, and from a global perspective, chromatin is organized in two main structural domains, euchromatin and heterochromatin, which differ with respect to the degree of compaction. The majority of genes reside in the loosely packed, active euchromatin, in which they remain easily accessible. On the other hand, the far more condensed heterochromatin includes silenced elements such as repetitive sequences and transposons, and is also found in centromeric and telomeric chromosome regions (Martens et al., 2005; Schueler and Sullivan 2006; Blasco 2007; Slotkin and Martienssen 2007). The organization of DNA as chromatin is

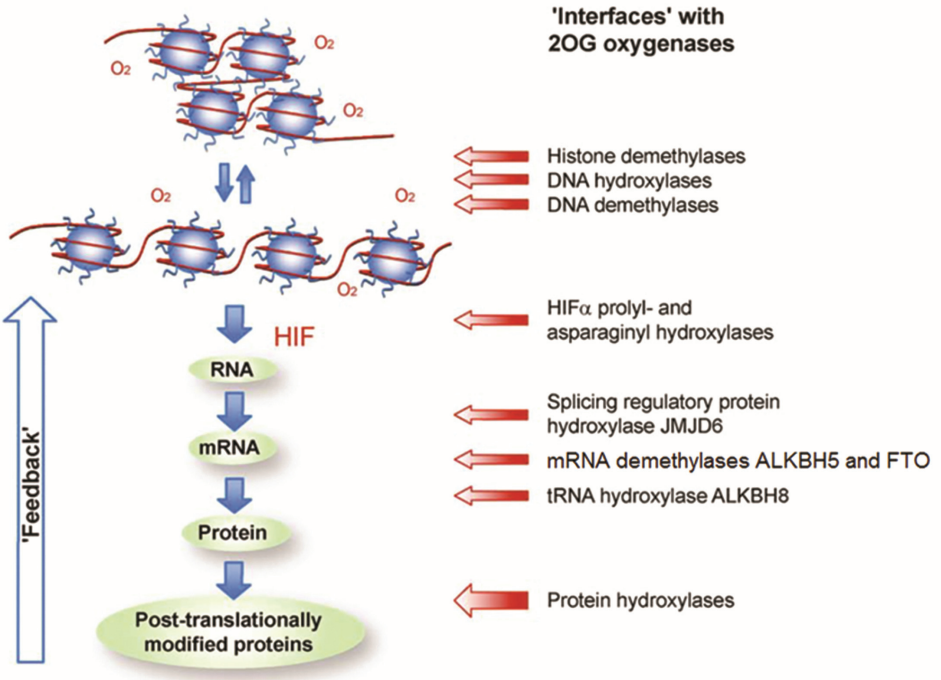


Figure 4. Proposed involvement of Fe(II)/2OG-dependent dioxygenases in regulation of gene expression. In addition to the JmjC-domain family of histone demethylases, DNA hydroxylases/demethylases in the TET family might also contribute to regulation at the chromatin level. While transcription is regulated in an oxygen-dependent manner through hydroxylation of the transcription factor HIF by FIH and PHD enzymes, post-transcriptional regulation might be performed by the mRNA demethylases ALKBH5 and FTO. Moreover, the tRNA hydroxylase ALKBH8 may function in regulation of translation, and proteins may be regulated post-translationally through modification by protein hydroxylases such as prolyl-3-hydroxylase (P3H) and prolyl-4-hydroxylase (P4H). Adapted from Loenarz and Schofield (2011).

thus important to preserve genome integrity during cell division and to facilitate DNA-dependent processes such as replication, transcription and DNA repair. This requirement is accommodated by the dynamics of the highly ordered chromatin structure which is, as described below, regulated through modification of its individual components.

DNA modifications

DNA is subjected to methylation at the 5-position in cytosine, resulting in 5-methylcytosine (5-meC). The covalent attachment of a methyl group in this position is

performed by enzymes in the DNA methyltransferase (DNMT) family. The genomic DNA methylation pattern is initially established by the *de novo* methyltransferases DNMT3A and DNMT3B during early embryogenesis (Okano et al., 1999), and further transmitted through subsequent cell generations by the maintenance methyltransferase DNMT1 which targets newly synthesized, hemimethylated DNA in replication foci (Leonhardt et al., 1992).

DNA methylation predominantly occurs in the context of CpG dinucleotide sequences. The CpG content of the vertebrate genome is asymmetrically distributed as a consequence of the mutagenic effect of 5-meC, which is frequently deaminated to thymine in a spontaneous manner (Coulondre et al., 1978; Bird 1980). Hence, CpG dinucleotides are predominantly clustered in CpG-rich regions, denoted CpG islands (CGIs), which are normally found in the unmethylated state (for a review, see Deaton and Bird (2011)). In the vertebrate genome, CGIs are associated with the promoters of ubiquitously expressed genes as well as many tissue-specific genes (Larsen et al., 1992; Zhu et al., 2008), and aberrant inactivation of such genes have been reported to correlate with methylation of their promoter CGIs in several cancers (Esteller 2007). Moreover, promoter CGI methylation is correlated with stable gene inactivation in processes such as X chromosome inactivation and imprinting (Edwards and Ferguson-Smith 2007; Reik 2007). Thus, 5-meC is generally viewed as a repressive epigenetic mark associated with long-term gene silencing.

One mechanism by which 5-meC functions to inhibit gene expression is through the ability to directly prevent transcription factors to bind their target DNA sequences (Iguchi-Ariga and Schaffner 1989; Prendergast and Ziff 1991). Additionally, this modification is recognized by methyl-CpG-binding proteins (MBPs) which further recruit chromatin remodeling factors and transcriptional repressors such as histone deacetylases (reviewed in Klose and Bird (2006)), thus also mediating gene inhibition in an indirect manner.

However, 5-meC is also present beyond promoter CpG islands and even in non-CpG sites (Lister et al., 2009). Intriguingly, methylation of CpG islands located in gene bodies is, in contradiction with the silencing function of promoter CpG methylation, correlated with active gene expression, a contradiction which is known as the DNA methylation paradox (Jones 1999). Although the function of gene body methylation is still not fully known, it has been proposed to be a mechanism for silencing of transposable elements, as methylation of their start sites result in suppression while transcription of the host gene in which they reside is not affected (Yoder et al., 1997). Regulation of alternative splicing has been suggested as another possible function, as exon-intron boundaries display a marked decrease in the degree of methylation (Laurent et al., 2010).

Another modified DNA base is β-D-glucopyranosyloxymethyluracil (base J), which is present in the genomes of certain protozoan parasites including trypanosomes, but has, notably, not been found in metazoans (van Leeuwen et al., 1998). Synthesis of this hypermodified base occurs in a two-step pathway in which the first step involves hydroxylation of thymidine in dsDNA by the Fe(II)/2OG-dependent dioxygenases J-binding protein (JBP) 1 and 2 (Yu et al., 2007; Cliffe et al., 2009; Vainio et al., 2009; Cliffe et al., 2012). The resulting intermediate, 5-hydroxymethyluracil (5-hmU), is subsequently glycosylated by an unknown glycosyltransferase to form base J (Cliffe et al., 2012). A gene regulatory function of base J was suggested with the finding that a subfraction localizes to genomic regions flanking RNA Polymerase II (Pol II) transcription sites (Cliffe et al., 2010). JBP1/2 deletion studies revealed that loss of base J coincided with decreased nucleosome density as well as increased histone acetylation and promoter occupancy by Pol II, thus suggesting an epigenetic mechanism for base J-mediated repression of transcription initiation (Ekanayake and Sabatini 2011; Ekanayake et al., 2011).

DNA methylation reversal

Although DNA methylation was traditionally considered a stable epigenetic mark, it is now broadly accepted that the two stages of epigenetic reprogramming that takes place during mammalian embryonic development involve an active mechanism for reversal of cytosine methylation, as suggested by the reported observations of rapid depletion of 5-meC in the paternal pronucleus of the zygote prior to the first mitosis, as well as a similar global 5-meC reduction during gametogenesis (Mayer et al., 2000; Oswald et al., 2000; Hajkova et al., 2002). Additionally, active demethylation has also been suggested to occur gene specifically in somatic cells (Bruniquel and Schwartz 2003). However, no demethylase with specificity for 5-meC has yet been discovered, despite an intense search for responsible enzymes.

The discovery that the human ten-eleven translocation 1 (TET1) protein catalyzes conversion of 5-meC to 5-hydroxymethylcytosine (5-hmC) was important for the understanding of the dynamics of the 5-meC modification (Figure 5A) (Tahiliani et al., 2009). The mammalian TET proteins (TET1, TET2 and TET3) compose a subfamily of the Fe(II)/2OG-dependent dioxygenases, and all three enzymes have been reported to possess 5-meC hydroxylase activity in presence of Fe(II) and 2-oxoglutarate (Tahiliani et al., 2009; Ito et al., 2010). Moreover, these proteins can further oxidize 5-hmC to 5-formylcytosine (5-fC) and 5-carboxylcytosine (5-caC, Figure 5A), thus generating two additional cytosine derivatives (He et al., 2011; Ito et al., 2011). As an attractive alternative to direct

demethylation of 5-meC, one could imagine an additional last step in which 5-caC is further decarboxylated to recover the unmodified cytosine. However, the existence of such a decarboxylase remains to be proven. Nonetheless, serial oxidation has indeed been suggested as a biologically important mechanism for active removal of 5-meC since the final product, 5-caC, is subjected to excision by thymine DNA glycosylase (TDG) which subsequently activates the base excision repair (BER) pathway (He et al., 2011; Maiti and Drohat 2011; Hashimoto et al., 2012). Moreover, 5-hmC has also been reported to be more prone to AID/APOBEC-mediated deamination than 5-meC, thus suggesting another possible multi-step mechanism for 5-meC demethylation in which the resulting 5-hydroxymethyluracil (5-hmU) is also a substrate for TDG (Cortellino et al., 2011; Hashimoto et al., 2012). Although the complete mechanism of active 5-meC demethylation is still not known, the present picture clearly suggests TET-mediated 5-meC oxidation to 5-hmC to be the first step.

The presence of 5-hmC in genomic DNA has proposed this modification to be a biologically relevant epigenetic mark rather than simply an intermediate in the 5-meC demethylation process (He et al., 2011; Ito et al., 2011; Pfaffeneder et al., 2011). Intriguingly, the previously reported decrease of 5-meC in the paternal pronucleus during zygotic reprogramming was recently shown to correlate with accumulation of 5-hmC, while the maternal pronucleus exhibited low 5-hmC and high 5-meC levels (Gu et al., 2011; Iqbal et al., 2011; Wossidlo et al., 2011). This conversion of 5-meC to 5-hmC was further shown to be catalyzed by TET3 (Gu et al., 2011). Notably, however, the paternal genome levels of 5-hmC have also been reported to be subjected to passive dilution upon replication (Inoue et al., 2011; Inoue and Zhang 2011). Moreover, it remains to be determined whether 5-hmC is recognized by reader proteins equivalent to the 5-meC binding MBPs. Clearly, additional effort is needed to determine the precise role of 5-hmC and its oxidation products in 5-meC demethylation as well as in the processes of chromatin regulation and epigenetic reprogramming.

Histone modifications

DNA is not the only chromatin component subjected to modifications. The N-terminal tails of the core histones, which protrude from the nucleosomes and are consequently available to modifying enzymes, display a wide variety of covalently linked chemical groups such as acetyl, methyl and phosphate (reviewed in Bannister and Kouzarides (2011)). The different modifications are introduced at the side chains of specific amino acid residues, including lysine (K), arginine (R) and serine (S), by histone modification enzymes including

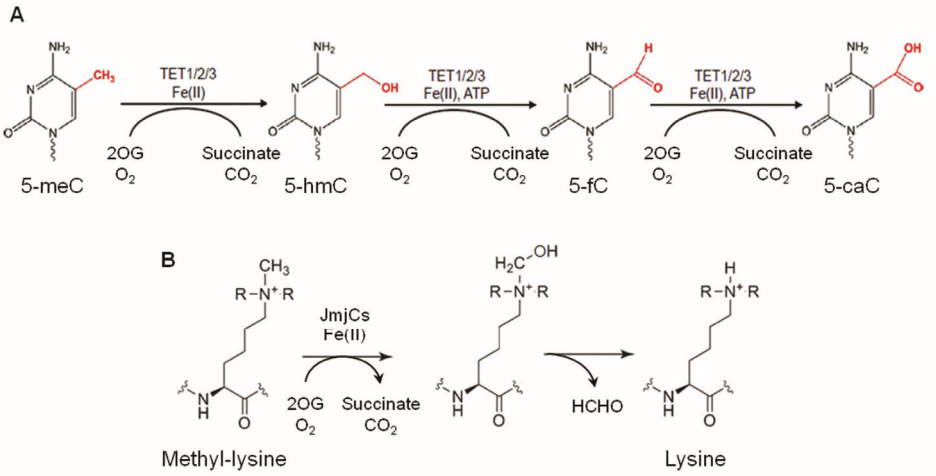


Figure 5. Oxidative reactions by chromatin-targeting Fe(II)/2OG-dependent dioxygenases. A) TET protein-mediated hydroxylation of 5-meC in DNA to 5-hmC. In presence of ATP the product can be further oxidized to the stable modifications 5-fC and 5-caC. **B)** JmjC-domain containing proteins catalyze demethylation of methyl-lysine in histones through hydroxylation of the methyl group(s), which is/are spontaneously released as formaldehyde (HCHO). Adapted from Loenarz and Schofield (2011) and Tan and Shi (2012).

histone acetyl- and methyltransferases (HATs and HMTs) as well as kinases, and the effects depend on the specific chemical group. Methylation differs from acetylation and phosphorylation in the way that the target arginine and lysine residues are not limited to modification with a single methyl group, but can acquire higher order methylation levels. While arginines can be dimethylated, either symmetrically or asymmetrically, lysines display the potential of both di- and trimethylation in addition to mono-methylation (reviewed in Bedford and Clarke (2009); Lan and Shi (2009)).

Generally, acetylation is associated with an open chromatin state as the positive charge of the affected lysine side chains are neutralized, leading to chromatin decondensation through decreased affinity of histones for DNA (Bannister and Kouzarides 2011). In a similar fashion, phosphorylation also provides increased chromatin access as addition of negative phosphate groups to serine residues result in a weakened histone-DNA interaction. In contrast, methylation does not influence on this electrostatic aspect of chromatin regulation as

the small, neutral methyl group does not alter histone charge and hence does not cause direct chromatin structure perturbation.

A second mechanism by which histone modifications regulate chromatin structure is by providing binding sites for effector proteins which function in ATP-dependent structure remodeling or recruitment of additional modification enzymes. Among the proteins that are recruited to methyl groups are the chromodomain and Tudor domain containing proteins and, similarly, acetyl and phosphate groups are recognized by effector protein modules such as bromodomains and 14-3-3 domains, respectively (reviewed in Yap and Zhou (2010)). Notably, certain modifications can also prevent binding of effector proteins (Fischle et al., 2005; Iberg et al., 2008). By such indirect means, the modification-embedded information in the histones can be interpreted by chromatin effectors whose action further modulate the chromatin structure positively or negatively to bring about the biological outcome.

Several modifications occur in close proximity on the histone tails, suggesting that the binding ability of an effector protein to its target will also be influenced by the modifications on neighboring residues. According to the "histone code" hypothesis, the different histone tail modifications act together, thus increasing the complexity of the regulatory system provided by histone modifications (Strahl and Allis 2000). This provides another dimension to the aspect of chromatin regulation mediated through post-translational histone marks, and increasing evidence for the existence of combinatorial effects of histone modifications is currently emerging (reviewed in Rando (2012)), as several effector proteins have been reported to possess binding modules that recognize a modification only in a certain context.

Completing the circle of histone modification regulation, enzymes with the ability to catalyze histone mark removal, like histone deacetylases (HDACs), demethylases (HDMs) and phosphatases, exist to counteract the activity of those introducing the modifications (reviewed in Bannister and Kouzarides (2011)). Thus, the antagonistic actions of histone code writers and erasers thus function to control chromatin structure dynamics in order to regulate transcription and other chromatin-dependent processes such as replication and DNA repair.

Histone methylation reversal

In contrast to other histone modifications, methylation was considered a static and irreversible mark as no enzyme with histone demethylase activity was known. Nearly a decade ago, this view was radically changed with the identification of lysine-specific demethylase 1 (LSD1), an amine oxidase with the ability to catalyze reversal of lysine methylation (Shi et al., 2004). However, this chemical mechanism requires two unpaired

electrons in the nitrogen on which the methyl groups are situated, thus restricting the LSD1 activity to mono- and di-methylated lysines (Shi et al., 2004).

Interestingly, histone lysine demethylase activity was demonstrated also for a subgroup of the Jumonji C (JmjC) domain-containing enzymes, referred to as the JmjC domain-contaig histone demethylases (JHDMs) (Cloos et al., 2006; Fodor et al., 2006; Klose et al., 2006; Tsukada et al., 2006; Whetstine et al., 2006; Yamane et al., 2006; Hong et al., 2007; Qi et al., 2010). Belonging to the superfamily of Fe(II)/2OG-dependent dioxygenases, the JmjC proteins employ the same demethylation mechanism used by AlkB to repair methylation damage in nucleic acids, i. e. hydroxylation of the histone lysine methyl group which is consequently destabilized and spontaneously released as formaldehyde (Tsukada et al., 2006). This reaction mechanism is, in contrast to that of the LSD1 family, compatible with reversal of mono-, di- and trimethyl-lysine (Figure 5B). However, as described in more detail below, the initially reported histone lysine and arginine demethylase activities of the JmjC-domain enzymes Jmjd5 and Jmjd6, respectively, have subsequently been challenged.

The mammalian ALKBH family might also be involved in histone demethylation, as ALKBH1 has been proposed to regulate gene expression through its histone H2A dioxygenase activity (Lando et al., 2012; Ougland et al., 2012). Moreover, the *Schizosaccharomyces pombe* (*S. pombe*) AlkB family member Ofd2, which is highly similar to ALKBH1, was reported to regulate the expression of oxidative phosphorylation genes during hypoxia through oxidation of histone H2A (Lando et al., 2012). However, this activity was suggested to result in H2A hydroxylation, not demethylation (Lando et al., 2012).

Evidently, the dynamics of histone methylation are not yet fully understood. The present picture of its regulation leaves lysine 79 in histone H3 (H3K79) as the only known methylation-subjected histone lysine residue for which no demethylase has been identified (Lan and Shi 2009; Liu et al., 2010; Qi et al., 2010; Stender et al., 2012). Furthermore, the debate concerning the existence of direct methyl-arginine reversal persists, as the only enzyme known to antagonize the action of the methyl-arginine generating protein arginine methyltransferases (PRMTs), peptidyl arginine deiminase 4 (PADI4), converts methyl-arginine to the non-standard amino acid citrulline through deimination, rather than catalyzing its demethylation to arginine (Cuthbert et al., 2004; Wang et al., 2004). However, the mechanism used by Fe(II)/2OG-dependent dioxygenases could easily be imagined to catalyze reversal of methylarginine as well as methylated H3K79. Hence, if such enzyme activities exist, the uncharacterized members of this protein family represent good candidates.

Protein hydroxylation

Gene regulation by Fe(II)/2OG-dependent dioxygenases also involves protein hydroxylation. An example of such a hydroxylated protein is the transcription factor hypoxia-inducible factor (HIF), which functions to stimulate gene expression through binding to hypoxia response elements (HREs) in the promoters of a set of genes involved in the cellular response to hypoxia (Semenza 1999; Kaelin and Ratcliffe 2008). HIF is composed of two subunits, HIF-1α and HIF-1β which upon hypoxia dimerize to induce transcription of target genes (Wang and Semenza 1995). However, under normoxic conditions HIF-1α is negatively regulated by two pathways which both involve the activity of Fe(II)/2OG-dependent dioxygenases, thereby inhibiting target gene expression. Engaged in the first pathway is the asparaginyl hydroxylase factor inhibiting HIF (FIH), which through hydroxylation of Asn803 in the C-terminal transactivation domain of HIF-1α represses the transcriptional activity of HIF by preventing recruitment of the co-activator p300 (Lando et al., 2002). The other pathway involves prolyl hydroxylase domain protein (PHD)-mediated hydroxylation of two proline residues (Pro402 and Pro564) in HIF-1α, leading to its ubiquitination and proteasomal degradation (Ivan et al., 2001; Jaakkola et al., 2001).

Another transcription factor, nuclear factor of activated T-cells calcineurin-dependent 1 (NFATc1), was recently reported to be regulated through lysine hydroxylation by the JmjC-domain containing enzyme Jmjd5 (Youn et al., 2012). The previously reported histone lysine demethylase activity (Hsia et al., 2010) was not detected, suggesting protein hydroxylation as the biologically relevant Jmjd5 activity. Similarly, Jmjd6, initially identified as the first histone arginine demethylase (Chang et al., 2007), was later found to not target histone methyl-arginine, but rather possess lysine hydroxylase activity towards the RNA splicing factor U2AF65 (Webby et al., 2009). Supporting the suggested functions of Jmjd5 and Jmjd6 in protein hydroxylation, these proteins were on the basis of structural studies shown to be homologous to the asparaginyl hydroxylase FIH (Elkins et al., 2003; Del Rizzo et al., 2012).

RNA modifications

Traditionally, post-transcriptional RNA modifications were viewed to play a role in fine-tuning of RNA structure and function (Bokar 2005). However, following the discovery that some of these modifications are reversible (Yi and Pan 2011), it was suggested that RNA modifications might function in regulation of gene expression, analogous to the dynamic methylations found in DNA and histones (He 2010). One of the well-known RNA modifications recently found to be reversible, N^6-methyladenosine (6-meA) in mRNA, has

previously been proposed to affect translation efficiency, since the methylated dihydrofolate reductase (DHFR) transcript was found to be translated more efficiently than the non-methylated one (Tuck et al., 1999). Further supporting a gene regulatory role of 6-meA and FTO, the Fe(II)/2OG-dependent dioxygenase responsible for its demethylation (Jia et al., 2011), 6-meA was reported to be important for pre-mRNA splicing (Dominissini et al., 2012). Very recently, a second Fe(II)/2OG-dependent dioxygenase, ALKBH5, was demonstrated to reverse 6-meA in mRNA (Zheng et al., 2012). ALKBH5 demethylase activity was further shown to be important for proper mRNA export to the cytoplasm, suggesting that 6-meA functions in post-transcriptional regulation of gene expression at the level of mRNA (Zheng et al., 2012). In addition, another RNA modification, mchm^{5}U, which is introduced in the anticodon-loop wobble position of tRNA^{Gly} by ALKBH8-mediated hydroxylation of mcm^{5}U (Fu et al., 2010; van den Born et al., 2011), has also been speculated to function in gene regulation since the modification alters the affinity of tRNA to specific mRNA codons (Fu et al., 2010).

Finally, structural studies of Jmjd6 suggested ssRNA, but not ssDNA or double-stranded nucleic acids, to be a biologically relevant substrate of this enzyme (Hong et al., 2010). Whether this RNA binding is related to the catalytic activity of Jmjd6 remains to be established. However, given the previously shown association with the mRNA splicing factor U2AF65 (Webby et al., 2009), Jmjd6 was further proposed to possibly function in regulation of alternative splicing through pre-mRNA demethylation or hydroxylation (Hong et al., 2010).

Thus, post-transcriptional regulation of gene expression at the level of RNA appears to be a biologically relevant process in which Fe(II)/2OG-dependent dioxygenases play a central role.

Aim of study

The *E. coli* AlkB protein is a Fe(II) and 2-oxoglutarate-dependent DNA repair enzyme involved in reversal of alkylation damage (Aravind and Koonin 2001; Falnes et al., 2002; Trewick et al., 2002). Initially, a single human homolog was identified (Wei et al., 1996). However, two subsequent studies revealed, through bioinformatics, the presence of six additional homologs (Kurowski et al., 2003; Gerken et al., 2007), and to date this protein family comprises nine proteins, ALKBH1-8 and the fat mass and obesity-related protein (FTO). While ALKBH2 and ALKBH3 were shown to possess similar functions as AlkB relatively soon after their discovery (Aas et al., 2003), the functions of the remaining ALKBH proteins have just recently started to emerge. Initially, the large number of homologs led to the correct suggestion that some of them does not function in DNA repair.

The aim of this study has been to improve our knowledge on the functions of uncharacterized members of this mammalian protein family. While the main focus has been characterization of ALKBH4, we have also touched upon ALKBH7. The initial approach chosen to investigate the functions of these proteins involved Yeast two-hybrid screens to identify interaction partners and, thereby, biological processes in which these proteins are involved (Paper II). Moreover, based on our finding that ALKBH4 possesses a unique cysteine-rich cluster in the N-terminus (Paper I), we aimed at obtaining insights into the iron-binding properties of ALKBH4. This issue was addressed through site-directed mutagenesis of residues in the cysteine cluster, as well as in the previously predicted active site iron-binding motif, and subsequent spectroscopic studies and measurements of the *in vitro* enzyme activity of the recombinant proteins.

Although the activities of some mammalian ALKBH proteins have been disclosed during the time period of this study, the biological function of several of these enzymes, including ALKBH4 and ALKBH7, still remains to be revealed. The work underlying this thesis is thus part of the continuing project with the objective of elucidating the functions of the mammalian ALKBH proteins.

Summary of papers

Paper I

Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4

Bjørnstad L. G., Zoppellaro G., Tomter A. B., Falnes P. Ø. and Andersson K. K.

Biochem. J. (2011) 434, 391-398

In this first report on ALKBH4, we describe the spectroscopic properties of the predicted iron-binding motif in the active site of this protein through electron paramagnetic resonance (EPR) and ultraviolet-visible (UV-vis) spectroscopy. Site-directed mutagenesis of the two proximal residues in the HXD…H (H169A-D171A-H254) triad predicted to bind iron confirmed this motif to indeed be essential for iron coordination.

In addition, we reported the presence of a unique conserved motif comprising four cysteine residues in the N-terminus of putative ALKBH4 orthologs. However, our speculation of whether this motif composes an iron-sulfur cluster was not supported as abrogation of the cysteine cluster did not affect the iron-binding features of ALKBH4, suggesting that ALKBH4 is a mononuclear iron protein.

Finally, we demonstrated that recombinant ALKBH4 is able to perform decarboxylation of the co-substrate 2-oxoglutarate (2OG) in absence of primary substrate, an uncoupled reaction typical for Fe(II)/2OG-dependent dioxygenases. Consistent with the results obtained from the spectroscopic analysis, both ferrous iron (Fe(II)) and an intact iron-binding site, were required for such activity, while the cysteine motif was not. In conclusion, we provided experimental data supporting that ALKBH4 is an active enzyme and a *bona fide* member of the superfamily of Fe(II)/2-OG-dependent dioxygenases.

Paper II

Human ALKBH4 interacts with proteins associated with transcription

Bjørnstad L. G., Meza T. J., Otterlei M., Olafsrud S. M., Meza-Zepeda L. A. and Falnes P. Ø.

PLoS ONE (2012) 7(11): e490945. doi:10.1371/journal.pone.0049045

This study focused on increasing our knowledge about the functions of ALKBH4 and the non-characterized paralog ALKBH7. In order to identify biological processes in which the two proteins exert their functions, we searched for interaction partners of the two proteins through yeast two-hybrid screens. While this approach did not reveal any binding partners for ALKBH7, we found ALKBH4 to interact with several proteins associated with transcription. Furthermore, these interactions were all indicated to be mediated through chromatin-associated domains, suggesting involvement of ALKBH4 in chromatin regulation. Moreover, the interactions were not dependent on the enzymatic activity of ALKBH4.

Supporting the results from the yeast two-hybrid analysis, three of the transcription-associated proteins showed nuclear co-localization with ALKBH4 in nucleoplasm and nucleoli. Interestingly, we also noted these proteins to exist as fusion partners of the histone methyltransferase MLL in the acute leukemia subtype mixed-lineage leukemia.

Furthermore, to address the possible effect of ALKBH4 and ALKBH7 on transcriptional regulation, we performed a comparison of the gene expression profiles of cells over-expressing ALKBH4 or ALKBH7 with those of the respective non-overexpressing cells. While ALKBH4 over-expression had a remarkably small effect on the global gene expression pattern, indicating no role in transcription regulation at the genome-wide level, over-expression of ALKBH7 resulted in larger perturbations in the gene expression pattern. Gene ontology (GO) analysis of the genes affected by ectopic ALKBH7 expression revealed enrichment of genes involved in processes such as cell cycle regulation, DNA damage and spermatogenesis. Consistent with this, we found three genes associated with meiotic recombination among the twenty most up-regulated genes. Additionally, a CpG methylation profile analysis excluded involvement of both proteins in regulation of global CpG methylation levels.

Taken together, the results obtained in this study suggest some processes in which ALKBH4 and ALKBH7 might function and provide a base for further investigation of the biological and biochemical functions of these proteins.

Discussion

The DNA repair enzyme AlkB from *E. coli* is a Fe(II)/2OG-dependent dioxygenase for which nine mammalian homologs, ALKBH1-8 and FTO, have been identified through bioinformatics. The study underlying this thesis has contributed to increased insight into the functions of two uncharacterized human ALKBH proteins, ALKBH4 and ALKBH7. When this study was initiated, the biological function of only one of the mammalian proteins, the functional AlkB homolog ALKBH2, had been established, and it has proven difficult to elucidate the roles of the remaining family members. However, their functions have started to slowly emerge, revealing involvement in other cellular pathways, as was proposed already a decade ago. Our findings indicate that both ALKBH4 and ALKBH7 are likely involved in processes other than repair of alkylation lesions in nucleic acids and provide a basis for further research aiming to disclose the functions of these interesting enzymes.

Biological function of ALKBH4

ALKBH4 in gene regulation – transcription and chromatin

In an attempt to increase our knowledge on the biological function of ALKBH4, we used the yeast two-hybrid (Y2H) system to identify interaction partners (Paper II). Several of its putative partners are associated with transcriptional regulation, suggesting ALKBH4 to be involved in regulation of gene expression. As the regions found to mediate binding to ALKBH4 in all cases comprised chromatin-binding domains, regulation by ALKBH4 may occur at the chromatin level.

Two of the ALKBH4-interacting proteins identified in the Y2H system, the related proteins AF9 and ENL, are components of the super elongation complex (SEC) (Lin et al., 2010; Mohan et al., 2010; He et al., 2011), which regulates the elongation step of transcription (references in Smith et al. (2011)). SEC recruitment to chromatin-bound RNA Polymerase II occurs via the AF9/ENL YEATS domain to sequence-specific factors such as mixed-lineage leukemia (MLL) (Yokoyama et al., 2010), or to the general Polymerase-associated factor complex (PAFc) (He et al., 2011). The region of AF9/ENL required for the interaction with ALKBH4 was the YEATS domain. Further supporting this, ALKBH4 co-localized with the

ENL YEATS domain in HeLa cells, while a YEATS domain truncation of ENL did not (Paper II). Thus, ALKBH4 might function in regulation of SEC recruitment to target sites.

Furthermore, the transcriptional co-activator and histone acetyl transferase (HAT) p300, here found to interact with ALKBH4 (Paper II), interacts with the non-phosphorylated, transcription initiation-associated form of RNA polymerase II (Cho et al., 1998). The p300 region responsible for the interaction to ALKBH4, the bromodomain (BRD) and plant homeodomain (PHD), has been demonstrated to function in recruitment of the protein to acetylated nucleosomes (Ragvin et al., 2004).

Moreover, we detected the heat-shock transcription factor HSF4 to interact with ALKBH4 through its DNA binding domain (DBD) (Paper II). Supporting this interaction, HSF4 targets the Hsp70-encoding *HSPA1* gene (Tu et al., 2006), which was upregulated upon ALKBH4 over-expression (Paper II). Moreover, *HSPA8,* the gene encoding the constitutively expressed heat-shock protein Hsc70, was also upregulated. This further strengthens the suggested involvement of ALKBH4 in regulation of transcription elongation, since SEC components were previously reported to localize to *HSPA1* upon heat-shock (Lin et al., 2010).

In conclusion, our findings suggest that ALKBH4 functions in gene regulation at the level of transcription, possibly by regulating recruitment of transcription complexes to chromatin target sites. However, it remains to be determined whether ALKBH4 affects transcription in a positive or negative manner, and the molecular mechanism for the suggested regulatory function of ALKBH4 also requires further investigation.

ALKBH4 is not involved in regulation of 5-methylcytosine

Transcriptionally repressed genes are generally associated with presence of 5-methylcytosine (5-meC) in their promoters. The chemical mechanism by which AlkB, ALKBH2 and ALKBH3 reverse nucleic acid alkylation damage could be envisioned to be utilized for oxidation of a methyl group in the 5-position of cytosine, and hydroxylation of 5-meC was proposed as a potential function for uncharacterized mammalian AlkB homologs (Sedgwick et al., 2007). When this study was initiated, it had not yet been shown that 5-meC hydroxylation is actually the reaction catalyzed by the TET proteins. We thus speculated whether ALKBH4, being an Fe(II)/2OG-dependent dioxygenase potentially involved in regulation of transcription, could possess similar activity. However, consistent with the subsequent identification of TET1-mediated conversion of 5-meC to 5-hydroxymethylcytosine (5-hmC) (Tahiliani et al., 2009), we observed no effect on the global 5-meC pattern upon over-expression of ALKBH4 (Paper II).

Mixed-lineage leukemia – a role for ALKBH4?

Mixed-lineage leukemia (MLL) is a particularly aggressive subtype of acute leukemia which is associated with chromosomal translocations involving the *MLL* gene. MLL is a histone methyltransferase with specificity for lysine 4 in histone H3 (H3K4) (Milne et al., 2002). However, leukemic rearrangements result in chimeric proteins in which the N-terminal of MLL is fused in-frame to the C-terminal part of the fusion partner, and loss of H3K4 methyltransferase activity. A large number of translocation partners has been reported (Meyer et al., 2009), and their structural and functional diversity has contributed to confusion with respect to the molecular mechanism by which the MLL-fusions cause leukemia. MLL leukemias generally display aberrant H3K79 dimethylation profiles and dysregulated *Hox* gene expression (Armstrong et al., 2002; Ferrando et al., 2003; Krivtsov et al., 2008). Hence, the underlying mechanism for transformation was believed to occur through misrecruitment of the H3K79 methyltransferase DOT1L to MLL target genes by the fusion partner. However, recent biochemical studies revealed that the most common fusion partners together exist in the SEC (Lin et al., 2010; Yokoyama et al., 2010). The predominant mechanism for MLL-fusion mediated leukemogenesis was therefore suggested to be constitutive recruitment of SEC, rather than DOT1L, to MLL target genes, resulting in their aberrant activation through circumvention of the transcriptional initiation-to-elongation checkpoint (Mohan et al., 2010). The biological significance of increased H3K79 methylation in leukemic cells is still not clear, although the DOT1L complex is essential for transformation (Nguyen et al., 2011).

Intriguingly, three of the ALKBH4 partners identified in this study, AF9, ENL and p300 (Paper II), are fused to MLL in leukemia (Iida et al., 1993; Rubnitz et al., 1994; Ida et al., 1997). Both AF9 and ENL interact with DOT1L and SEC in a mutually exclusive manner (Lin et al., 2010; Yokoyama et al., 2010; Biswas et al., 2011; He et al., 2011), and the corresponding MLL-fusions are associated with constitutive recruitment of SEC and DOT1L to MLL target genes, suggesting that the resulting leukemic transformation is caused by persistent RNA Polymerase II-mediated gene transcription and H3K79 hypermethylation (Figure 6A) (Yokoyama et al., 2010; Bernt et al., 2011; Biswas et al., 2011). In contrast, the less frequent MLL-p300/CBP fusions are thought to promote leukemogenesis through aberrant histone acetylation, since the fusion proteins retain the histone acetyl transferase (HAT) domain (Figure 6B) (Ida et al., 1997; Slany 2009). The use of different mechanisms by MLL-AF9/ENL and MLL-p300 fusion proteins makes it difficult to suggest a specific role for ALKBH4 in MLL-based transformation. Direct participation of ALKBH4 in MLL-p300

mediated leukemogenesis could readily be imagined since the region of p300 necessary and sufficient for leukemic transformation comprises, in addition to the HAT domain, the bromodomain (BRD) and the PHD finger (Ida et al., 1997), both of which we found to be involved in mediating the interaction with ALKBH4 (Paper II). However, the ALKBH4-interacting YEATS domain is typically not present in the fusion proteins MLL-AF9 and MLL-ENL (Rubnitz et al., 1994), suggesting that ALKBH4 may be indirectly linked to leukemogenesis through its impaired recruitment to AF9/ENL. Alternatively, since the physiological SEC and DOT1L complexes are both required for leukemogenesis (Yokoyama et al., 2010; Nguyen et al., 2011), a direct role of ALKBH4 in the transformation process might also be envisioned, given that ALKBH4 associates with at least one of the complexes.

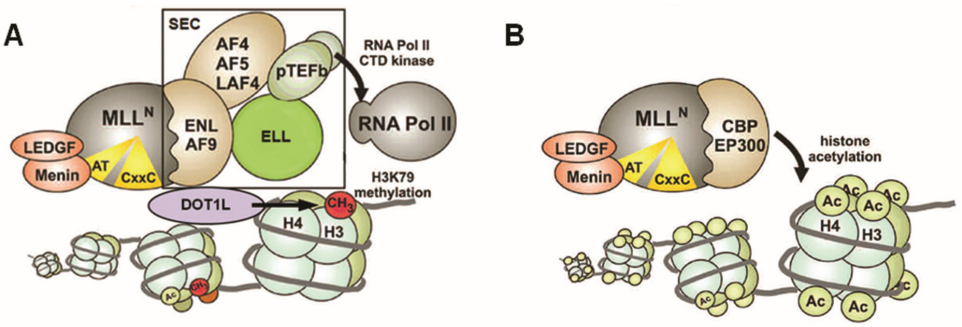


Figure 6. Schematic representation of the proposed mechanisms for leukemogenesis by the MLL-fusion proteins A) MLL-AF9/ENL and B) MLL-p300/CBP. In all cases, the N-terminal part of MLL (MLL^N) remains unchanged and retains its interaction with the chromatin-targeting factors Menin and lens-epithelium derived growth factor (LEDGF). However, the C-terminal part, including the SET methyltransferase domain, is exchanged with the fusion partner which is believed to be actively involved in leukemic transformation. MLL-AF9/ENL fusions are thought to promote leukemogenesis through stimulation of transcriptional elongation via the two elongation factors positive transcription elongation factor b (pTEFb) and eleven-nineteen lysine-rich leukemia (ELL), which both are components of the super elongation complex (SEC). Histone methylation by another AF9/ENL partner, the histone H3K79 methyltransferase DOT1L, is also associated with leukemic transformation. The mechanism used by MLL-p300/CBP fusions, which retain the histone acetyltransferase (HAT) activity of the fusion partners, is believed to be histone acetylation. Adapted from Slany (2009).

ALKBH4 in development

Gene targeting studies in mice have shown that both Af9 and Enl function in embryonic development (Collins et al., 2002; Doty et al., 2002). While *Enl*-inactivation leads to early embryonic lethality in mice (Doty et al., 2002), disruption of *Af9* is associated with perinatal lethality and altered expression of *Hox* genes, consistent with a function for AF9 in embryo patterning and antero-posterior axis formation (Collins et al., 2002). Initial *Hox* gene activation involves fibroblast growth factors (FGFs), and establishment of *Hox* expression patterns are dependent on retinoic acid (RA) signaling (reviewed in Deschamps et al. (1999)). Subsequently, the *Hox* expression domains established along the primary body axis are maintained by Trithorax and Polycomb group (TrxG and PcG) proteins which regulate the chromatin structure of the *Hox* gene cluster in a positive and negative manner, respectively (Gould 1997; Deschamps et al., 1999). Further supporting a role for AF9/ENL in regulation of *Hox* gene expression, both proteins have been reported to interact with the human PcG protein hPc3 (Garcia-Cuellar et al., 2001; Hemenway et al., 2001). Functions consistent with embryo patterning have also been reported for other herein-identified ALKBH4 partners. The homeotic transcription factor ATBF1 is expressed upon RA-induced neuronal differentiation (Ido et al., 1994) and HSF4 controls FGF repression during lens cell differentiation (Fujimoto et al., 2004). Likewise, the *Xenopus* ortholog of the non-transcription-associated ALKBH4 partner TES has also been implicated in regulation of axial elongation through FGF signaling (Dingwell and Smith 2006). Furthermore, studies in mice have shown that p300 is essential for embryogenesis (Yao et al., 1998), during which it is normally involved in neural tube and craniofacial development (Bhattacherjee et al., 2009). Notably, *Alkbh4*-targeted mice are early embryonic lethal (Nilsen and Klungland, unpublished data), demonstrating an essential developmental function also of ALKBH4. Taken together, this putative developmental function of ALKBH4 might possibly manifest through regulation of genes involved in developmental axis formation.

Biochemical function of ALKBH4

The biochemical functions of several mammalian ALKBH proteins, including ALKBH4, have not yet been determined. However, a previous study demonstrated that ALKBH4 is not engaged in DNA repair through demethylation of 1-meA or 3-meC, since the protein neither complemented the sensitivity of an *AlkB*-deficient *E. coli* strain towards transformation of MMS-treated single-stranded M13mp18 bacteriophage DNA, nor displayed

in vitro activity towards single- or double-stranded oligodeoxyribonucleotide substrates containing 1-meA or 3-meC (Bekkelund and Falnes, unpublished data; Lee et al. (2005)). The initial classification of ALKBH4 as a member of the mammalian AlkB family was based on a bioinformatics study (Kurowski et al., 2003), and has not been experimentally confirmed. Here we show, through spectroscopic studies, that ALKBH4 exhibits similar features as AlkB with respect to Fe(II) binding (Paper I), supporting the prediction of ALKBH4 being a member of the AlkB family of Fe(II)/2OG-dependent dioxygenases. Further supporting this, we demonstrated that ALKBH4 in presence of Fe(II) is able to decarboxylate 2-oxoglutarate, resulting in formation of succinate (Paper I), thus strongly suggesting that ALKBH4 has enzymatic oxygenase activity.

Hence, as an Fe(II)/2OG-dependent enzyme, ALKBH4 presumably catalyzes hydroxylation or demethylation of a yet unidentified protein or nucleic acid substrate, an issue that will be discussed in closer detail below.

Hydroxylase or demethylase activity?

Fe(II)/2OG-dependent dioxygenases typically catalyze hydroxylation reactions. While plants and microorganisms possess Fe(II)/2OG dioxygenases which catalyze oxidative reactions apart from the prevalent hydroxylations (Valegard et al., 1998; Turnbull et al., 2004; Vaillancourt et al., 2005), only hydroxylations have been reported for these enzymes in animals. As the first family member, *E. coli* AlkB was demonstrated to possess demethylase activity, in which methylation reversal occurs through hydroxylation of the target methyl group (Falnes et al., 2002; Trewick et al., 2002). Similarly, most of the characterized mammalian ALKBH proteins catalyze substrate demethylation via hydroxylation (Aas et al., 2003; Jia et al., 2011), however, conventional substrate hydroxylation is also represented among the ALKBHs, as target hydroxyl modification is performed by the oxygenase domain of ALKBH8 (Fu et al., 2010; van den Born et al., 2011). Thus, ALKBH4 probably possesses hydroxylase or hydroxylation-mediated demethylase activity.

Auto-hydroxylation of ALKBH4?

Spectroscopic studies have indicated several Fe(II)/2OG-dependent dioxygenases, including TfdA (Liu et al., 2001), TauD (Ryle et al., 2003), FIH (Chen et al., 2008) as well as AlkB (Henshaw et al., 2004), to catalyze auto-hydroxylation of an active-site aromatic residue in absence of primary substrate. Such activity involves oxidative conversion of the ferrous (Fe(II)) iron to the ferric (Fe(III)) species which complexes with the hydroxylated side chain,

resulting in enzyme inactivation. Although the role of self-hydroxylation is unclear, the auto-hydroxylated forms of both TauD and AlkB have been detected *in vivo* (Henshaw et al., 2004; Grzyska et al., 2007), thus suggesting biological relevance. A proposed function is regulation of enzyme activity, possibly to protect against formation of deleterious oxygen radicals in absence of substrate (Liu et al., 2001; Hausinger 2004) or, in the case of FIH, to regulate the hypoxia response (Chen et al., 2008). Self-hydroxylation of an active site residue has also been observed for ALKBH3, however, in contrast to the previously reported cases, the target site was a leucine rather than an aromatic residue (Sundheim et al., 2006). Notably, corresponding oxidation was also detected in all the AlkB family proteins in which this leucine is conserved, ALKBH2, ALKBH6 and AlkB (Sundheim et al., 2006).

Upon oxidation of the ALKBH4/Fe(II)/2OG complex we detected an absorption band similar to the oxidation products of the corresponding TfdA (Liu et al., 2001) and FIH (Chen et al., 2008) complexes, which was reported to arise from hydroxylated Tyr73 and Trp296, respectively (Liu et al., 2001; Ryle et al., 2003). However, we observed a similar absorption band also in absence of 2OG, suggesting that the complex giving rise to the chromophore of oxidized ALKBH4 may not depend on 2OG. Instead, it might possibly origin from complexed iron and oxygen. Together with our observation that the uncoupled 2-oxoglutarate turnover was substantially lower for ALKBH4 compared to AlkB (Paper I), this may possibly suggest that no self-hydroxylation target is present in or close to the active site of ALKBH4. However, in presence of the oxygen analog NO, oxidation of the ALKBH4/Fe(II)/2OG complex gave rise to an absorption band closely resembling the one of oxidized AlkB/Fe(II)/2OG (in absence of NO). While the latter absorption feature of AlkB was demonstrated to arise from auto-hydroxylation of Trp178 (Henshaw et al., 2004), Fe-NO could possibly be responsible for the chromophore we observed for ALKBH4/Fe(II)/2OG/NO. Evidently, additional effort is required to determine whether ALKBH4 displays auto-hydroxylase activity.

Nucleic acid or protein substrate?

Although both lipids and small molecules have been reported to be oxidized by Fe(II)/2OG-dependent dioxygenases, the substrates of these proteins typically comprise proteins and nucleic acids (Hausinger 2004). Notably, the ALKBH proteins for which substrates have been identified all target nucleic acids: the repair activities of ALKBH2 and ALKBH3 are directed towards lesions in DNA and RNA, respectively (Aas et al., 2003; Falnes et al., 2004), FTO and ALKBH8 are both involved in post-transcriptional modification

of RNA (Jia et al., 2011; van den Born et al., 2011). This implies that DNA or RNA might be the substrate also of ALKBH4. Consistent with the predominant location of ALKBH4 in the nucleus (Tsujikawa et al., 2007), including nucleolar speckles (Paper II), DNA or nuclear RNA such as mRNA or rRNA might represent possible substrates of this protein, as do smaller RNA species such as micro RNA (miRNA), small nuclear RNA (snRNA) or small nucleolar RNA (snoRNA).

However, ALKBH4 was previously, together with ALKBH1 and ALKBH7, suggested to act on protein substrates because of their low pI-values which could interfere with interaction with a nucleic acid substrate (Sedgwick et al., 2007). Possible ALKBH4 protein substrates are the interaction partners, which could be subjected to hydroxylation or demethylation. However, information in the Phosphosite database (Hornbeck et al., 2012) on post-translational protein modifications indicate that none of the ALKBH4 targets are hydroxylated and only two, p300 and eukaryotic translation initiation factor 3C (EIF3C), are methylated, disfavoring a hypothesis in which the partners are the ALKBH4 substrates.

The hydroxylase activities of the JmjC-domain containing Fe(II)/2OG-dependent dioxygenases target methyl groups on histone lysine side chains, leading to reversal of histone methylation by a mechanism analogous to that of AlkB proteins (Tsukada et al., 2006). Interestingly, two of the JmjC proteins initially reported to be histone lysine and arginine demethylases respectively, Jmjd5 and Jmjd6, possibly function *in vivo* to hydroxylate non-histone proteins (Webby et al., 2009; Del Rizzo et al., 2012; Youn et al., 2012), or even ssRNA (Hong et al., 2010). Similarly, histone demethylase activity was recently proposed also for the AlkB protein family, as ALKBH1 was suggested, as the first member, to demethylate one or two C-terminal lysine residues in histone H2A (Ougland et al., 2012). This suggests that that the individual Fe(II)/2OG-dependent dioxygenase subfamilies may not be restricted to either protein or nucleic acid substrates.

The over-representation of chromatin-associated proteins among the ALKBH4 partners identified in this study could imply that ALKBH4 possesses a histone substrate. Two findings made us speculate whether ALKBH4 could possibly target lysine 79 in histone H3 (H3K79), the one methylation-subjected histone lysine residue for which no demethylase has been identified: First, ALKBH4 interacts with AF9/ENL, a component of the DOT1L methyltransferase complex which is responsible for methylation of H3K79 (Paper II). Second, ALKBH4 over-expression resulted in down-regulation of *LGALS1* (Paper II), a biomarker of MLL-rearranged leukemia whose upregulation correlates with increased H3K79 methylation (Juszczynski et al., 2010). However, as we did not observe any effect of ALKBH4 over-

expression on the global H3K79 methylation status in HEK293 cells (Paper II), the question whether H3K79 is an ALKBH4 substrate remains open. The oxidation mechanism employed by Fe(II)/2OG-dependent dioxygenases is also compatible with arginine demethylation (Chang et al., 2007). As mentioned above, it is debated whether this proposed activity of Jmjd6 is biologically relevant, therefore direct reversal of arginine methylation *in vivo* has still to be demonstrated. However, one could imagine ALKBH4, being a Fe(II)/2OG-dependent dioxygenase, to possess a methylated histone, or non-histone, arginine as substrate.

Possible function of the ALKBH4 cysteine cluster

From the alignment of putative ALKBH4 orthologs from different metazoans, we observed a conserved motif comprising four cysteine residues in the N-terminus (Paper I). The high degree of conservation between species, together with absence of the motif in AlkB as well as in ALKBH proteins, implies its importance for ALKBH4 function. Abrogation of the cluster through site-directed mutagenesis of the cysteines to alanine had no effect on the uncoupled 2-oxoglutarate decarboxylase activity of ALKBH4 (Paper I), suggesting that the motif is not required for the basic enzymatic activity of ALKBH4. A BLAST search in the NCBI protein sequence database (Altschul et al., 1997), using the conserved cysteine-rich motif as input sequence, revealed no similarity to currently annotated domains. However, since cysteine-rich protein modules are often involved in metal coordination, the ALKBH4 cluster can be expected to constitute a metal site. Our spectroscopic analysis demonstrated that ALKBH4 contains a mononuclear iron site which comprises the HXD…H motif, suggesting that the cysteine motif is not an iron-sulfur (Fe-S) cluster (Paper I). However, it should be noted that an elaborate biosynthetic machinery is required for the formation of many Fe-S clusters, which therefore cannot assemble spontaneously *in vitro* upon reconstitution with iron (reviewed in Xu and Moller (2011)).

Zinc finger domains compose another class of metal-coordinating folds, which are remarkably diverse both in structure and function, and can bind both nucleic acids and proteins, as well as lipids (Laity et al., 2001). Zinc fingers are common in gene regulatory proteins and, given the possible role of ALKBH4 in chromatin regulation (Paper II), the N-terminal cysteine motif could likely represent a zinc-binding site. In support of such a hypothesis, parts of the ALKBH4 cysteine motif show similarities with known zinc fingers of the classical C_2HC- and C_2H_2-types, like those found in the nucleolar protein pNO40 and the transcriptional repressor ZBRK1 (Zheng et al., 2000; Chang et al., 2003). Notably, the zinc

fingers of ZBRK1 have been reported to bind DNA (Zheng et al., 2000), suggesting that also the ALKBH4 motif might function as a DNA or RNA recognition motif. However, single zinc fingers often mediate protein-protein interactions (Matthews and Sunde 2002). Thus, we speculated whether the ALKBH4 cysteine cluster is involved in interaction partner binding. However, preliminary data show that cysteine motif abrogation does not affect the co-localization pattern of ALKBH4 and ENL (Otterlei, Bjørnstad and Falnes, unpublished data), suggesting that the module might display other functions, possibly substrate binding. Alternatively, although less likely, the motif may simply have a structural purpose, like the Zn(II) coordinating cysteine cluster separating the methyltransferase and oxygenase domains of ALKBH8 (Pastore et al., 2012).

ALKBH7 – a possible function in DNA double strand break repair during meiotic recombination

In addition to the focus on characterization of ALKBH4, this study has also touched upon another ALKBH protein, ALKBH7. The molecular function of ALKBH7 is, like that of ALKBH4, unknown. However, the previous NCBI annotation of ALKBH7 as spermatogenesis-associated protein 11 (SPATA11) or "spermatogenesis cell proliferation-related protein" indicates ALKBH7 to function in spermatogenesis.

Through functional annotation analysis of genes affected by ALKBH7 over-expression, we identified spermatogenesis as one of the biological pathways enriched with differentially expressed genes, which is consistent with involvement of ALKBH7 in this process (Paper II). Furthermore, three of the twenty most upregulated genes, the meiosis-specific recombinase *disrupted meiotic cDNA 1 (DMC1)* and the regulatory proteins *decreased sperm survival 1 (DSS1)* and *male-specific lethal 3-like 1 (MSL3L1)*, are involved in meiotic recombination (Habu et al., 1996; Kojic et al., 2003; Dray et al., 2006; Sharma et al., 2010), a central event in spermatogenesis. Meiotic recombination is initiated with formation of DNA double-strand breaks (DSBs) induced by the endonuclease SPO11 (Keeney et al., 1997; Keeney et al., 1999). The strand breaks are resolved through homologous recombination repair (HRR) in a programmed cellular response involving checkpoint activation, DNA damage signaling and chromatin rearrangements (Fernandez-Capetillo et al., 2003; Miles et al., 2010). Furthermore, as revealed through studies of mouse gametocytes deficient in central components of the meiotic DSB response machinery, such as *DMC1*, unrepaired DSBs result in increased levels of apoptosis (Yoshida et al., 1998). Since

"Cell cycle and its regulation" was the pathway in this study that was mostly enriched with genes affected by ectopic ALKBH7 levels, followed by the categories "DNA-damage response" and "Apoptosis" (Paper II), our findings are consistent with a role for ALKBH7 in regulation of the response to meiotic DNA double strand breaks. Further supporting this, mouse *Alkbh7* expression is highest at the pachytene stage of male meiosis (GEO accesssion GDS2390; (Namekawa et al., 2006), the stage at which recombination occurs.

Notably, *ALKBH7* expression is not restricted to testis (Tsujikawa et al., 2007). Moreover, most of the meiotic recombination-associated genes that were found to be dysregulated upon ALKBH7 over-expression in this study are also involved in general repair of DNA double strand breaks (Gudmundsdottir et al., 2004; Kristensen et al., 2010; Sharma et al., 2010). Thus, the suggested involvement of ALKBH7 in homologous recombination might extend beyond meiosis to also comprise the mitosis-associated, homologous recombination-mediated, cellular response to double strand breaks.

Our findings thus support the previous annotation of ALKBH7 as a spermatogenesis-related protein, and further indicate that ALKBH7 might be involved in regulation of double strand break repair by homologous recombination during meiosis. Additional efforts are needed to elucidate the precise role of ALKBH7 in this process, and to determine whether its function also applies to general homologous recombination repair of double strand breaks.

Future perspectives

The research underlying this thesis has increased our insight into the field of mammalian ALKBH proteins and points towards some biological processes in which ALKBH4 and ALKBH7 may be involved. However, the molecular functions of both proteins remain to be determined.

In order to elucidate the function of ALKBH4 it would be useful to identify the role of the N-terminal cysteine cluster. Efforts should be invested in determination of the metal binding abilities of the motif, in particular with respect to zinc ions, and it would also be interesting to determine whether it represents a structural or functional fold. Given the possibility that ALKBH4 might possess a nucleic acid substrate, it would also be beneficial to examine the ability of the protein to bind DNA and RNA, and the possible involvement of the cysteine motif in such interaction.

Furthermore, the role of ALKBH4 in transcription should be addressed more closely. In this respect, it is of interest to determine whether the interaction of ALKBH4 with AF9/ENL is associated with gene regulation through the super elongation or DOT1L complexes. Considering the striking lack of alterations in the global gene expression pattern in HEK 293 cells before and after ALKBH4 over-expression, a possible function of ALKBH4 at a specific developmental stage or in response to a particular stimulus is likely, and should be investigated. Moreover, given that ALKBH4 interacts with three MLL fusion partners, it would be of interest to address the role of ALKBH4 in MLL leukemia.

Finally, it would also be interesting to further investigate the possible regulatory function of ALKBH7 in spermatogenesis and meiotic recombination.

References

Aas, P. A., Otterlei, M., Falnes, P. O., Vagbo, C. B., Skorpen, F., et al. (2003). "Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA." Nature **421**(6925): 859-863.

Aik, W., McDonough, M. A., Thalhammer, A., Chowdhury, R. and Schofield, C. J. (2012). "Role of the jelly-roll fold in substrate binding by 2-oxoglutarate oxygenases." Curr Opin Struct Biol.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., et al. (1997). "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs." Nucleic Acids Res **25**(17): 3389-3402.

Aravind, L. and Koonin, E. V. (2001). "The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases." Genome Biol **2**(3): RESEARCH0007.

Armstrong, S. A., Staunton, J. E., Silverman, L. B., Pieters, R., den Boer, M. L., et al. (2002). "MLL translocations specify a distinct gene expression profile that distinguishes a unique leukemia." Nat Genet **30**(1): 41-47.

Bannister, A. J. and Kouzarides, T. (2011). "Regulation of chromatin by histone modifications." Cell Res **21**(3): 381-395.

Bedford, M. T. and Clarke, S. G. (2009). "Protein arginine methylation in mammals: who, what, and why." Mol Cell **33**(1): 1-13.

Bernt, K. M., Zhu, N., Sinha, A. U., Vempati, S., Faber, J., et al. (2011). "MLL-rearranged leukemia is dependent on aberrant H3K79 methylation by DOT1L." Cancer Cell **20**(1): 66-78.

Bhattacherjee, V., Horn, K. H., Singh, S., Webb, C. L., Pisano, M. M., et al. (2009). "CBP/p300 and associated transcriptional co-activators exhibit distinct expression patterns during murine craniofacial and neural tube development." Int J Dev Biol **53**(7): 1097-1104.

Bird, A. P. (1980). "DNA methylation and the frequency of CpG in animal DNA." Nucleic Acids Res **8**(7): 1499-1504.

Biswas, D., Milne, T. A., Basrur, V., Kim, J., Elenitoba-Johnson, K. S., et al. (2011). "Function of leukemogenic mixed lineage leukemia 1 (MLL) fusion proteins through distinct partner protein complexes." Proc Natl Acad Sci U S A **108**(38): 15751-15756.

Bjornstad, L. G., Zoppellaro, G., Tomter, A. B., Falnes, P. O. and Andersson, K. K. (2011). "Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4." Biochem J **434**(3): 391-398.

Blasco, M. A. (2007). "The epigenetic regulation of mammalian telomeres." Nat Rev Genet **8**(4): 299-309.

Bokar, J. (2005). "The biosynthesis and functional roles of methylated nucleosides in eukaryotic mRNA." Fine-tuning of RNA Functions by Modification and Editing: 141-177.

Bruniquel, D. and Schwartz, R. H. (2003). "Selective, stable demethylation of the interleukin-2 gene enhances transcription by an active process." Nat Immunol **4**(3): 235-240.

Chang, B., Chen, Y., Zhao, Y. and Bruick, R. K. (2007). "JMJD6 is a histone arginine demethylase." Science **318**(5849): 444-447.

Chang, W. L., Lee, D. C., Leu, S., Huang, Y. M., Lu, M. C., et al. (2003). "Molecular characterization of a novel nucleolar protein, pNO40." Biochem Biophys Res Commun **307**(3): 569-577.

Chen, B., Liu, H., Sun, X. and Yang, C. G. (2010). "Mechanistic insight into the recognition of single-stranded and double-stranded DNA substrates by ABH2 and ABH3." Mol Biosyst **6**(11): 2143-2149.

Chen, B. J., Carroll, P. and Samson, L. (1994). "The Escherichia coli AlkB protein protects human cells against alkylation-induced toxicity." J Bacteriol **176**(20): 6255-6261.

Chen, Y. H., Comeaux, L. M., Eyles, S. J. and Knapp, M. J. (2008). "Auto-hydroxylation of FIH-1: an Fe(ii), alpha-ketoglutarate-dependent human hypoxia sensor." Chem Commun (Camb)(39): 4768-4770.

Cho, H., Orphanides, G., Sun, X., Yang, X. J., Ogryzko, V., et al. (1998). "A human RNA polymerase II complex containing factors that modify chromatin structure." Mol Cell Biol **18**(9): 5355-5363.

Church, C., Moir, L., McMurray, F., Girard, C., Banks, G. T., et al. (2010). "Overexpression of Fto leads to increased food intake and results in obesity." Nat Genet **42**(12): 1086-1092.

Cliffe, L. J., Hirsch, G., Wang, J., Ekanayake, D., Bullard, W., et al. (2012). "JBP1 and JBP2 proteins are Fe2+/2-oxoglutarate-dependent dioxygenases regulating hydroxylation of thymidine residues in trypanosome DNA." J Biol Chem **287**(24): 19886-19895.

Cliffe, L. J., Kieft, R., Southern, T., Birkeland, S. R., Marshall, M., et al. (2009). "JBP1 and JBP2 are two distinct thymidine hydroxylases involved in J biosynthesis in genomic DNA of African trypanosomes." Nucleic Acids Res **37**(5): 1452-1462.

Cliffe, L. J., Siegel, T. N., Marshall, M., Cross, G. A. and Sabatini, R. (2010). "Two thymidine hydroxylases differentially regulate the formation of glucosylated DNA at regions flanking polymerase II polycistronic transcription units throughout the genome of Trypanosoma brucei." Nucleic Acids Res **38**(12): 3923-3935.

Clifton, I. J., Doan, L. X., Sleeman, M. C., Topf, M., Suzuki, H., et al. (2003). "Crystal structure of carbapenem synthase (CarC)." J Biol Chem **278**(23): 20843-20850.

Cloos, P. A., Christensen, J., Agger, K., Maiolica, A., Rappsilber, J., et al. (2006). "The putative oncogene GASC1 demethylates tri- and dimethylated lysine 9 on histone H3." Nature **442**(7100): 307-311.

Collins, E. C., Appert, A., Ariza-McNaughton, L., Pannell, R., Yamada, Y., et al. (2002). "Mouse Af9 is a controller of embryo patterning, like Mll, whose human homologue fuses with Af9 after chromosomal translocation in leukemia." Mol Cell Biol **22**(20): 7313-7324.

Cortellino, S., Xu, J., Sannai, M., Moore, R., Caretti, E., et al. (2011). "Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair." Cell **146**(1): 67-79.

Coulondre, C., Miller, J. H., Farabaugh, P. J. and Gilbert, W. (1978). "Molecular basis of base substitution hotspots in Escherichia coli." Nature **274**(5673): 775-780.

Cuthbert, G. L., Daujat, S., Snowden, A. W., Erdjument-Bromage, H., Hagiwara, T., et al. (2004). "Histone deimination antagonizes arginine methylation." Cell **118**(5): 545-553.

Dango, S., Mosammaparast, N., Sowa, M. E., Xiong, L. J., Wu, F., et al. (2011). "DNA unwinding by ASCC3 helicase is coupled to ALKBH3-dependent DNA alkylation repair and cancer cell proliferation." Mol Cell **44**(3): 373-384.

Deaton, A. M. and Bird, A. (2011). "CpG islands and the regulation of transcription." Genes Dev **25**(10): 1010-1022.

Del Rizzo, P. A., Krishnan, S. and Trievel, R. C. (2012). "Crystal Structure and Functional Analysis of JMJD5 Indicate an Alternate Specificity and Function." Mol Cell Biol **32**(19): 4044-4052.

Delaney, J. C. and Essigmann, J. M. (2004). "Mutagenesis, genotoxicity, and repair of 1-methyladenine, 3-alkylcytosines, 1-methylguanine, and 3-methylthymine in alkB Escherichia coli." Proc Natl Acad Sci U S A **101**(39): 14051-14056.

Delaney, J. C., Smeester, L., Wong, C., Frick, L. E., Taghizadeh, K., et al. (2005). "AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo." Nat Struct Mol Biol **12**(10): 855-860.

Demple, B., Sedgwick, B., Robins, P., Totty, N., Waterfield, M. D., et al. (1985). "Active site and complete sequence of the suicidal methyltransferase that counters alkylation mutagenesis." Proc Natl Acad Sci U S A **82**(9): 2688-2692.

Deschamps, J., van den Akker, E., Forlani, S., De Graaff, W., Oosterveen, T., et al. (1999). "Initiation, establishment and maintenance of Hox gene expression patterns in the mouse." Int J Dev Biol **43**(7): 635-650.

Dinglay, S., Trewick, S. C., Lindahl, T. and Sedgwick, B. (2000). "Defective processing of methylated single-stranded DNA by E. coli AlkB mutants." Genes Dev **14**(16): 2097-2105.

Dingwell, K. S. and Smith, J. C. (2006). "Tes regulates neural crest migration and axial elongation in Xenopus." Dev Biol **293**(1): 252-267.

Dominissini, D., Moshitch-Moshkovitz, S., Schwartz, S., Salmon-Divon, M., Ungar, L., et al. (2012). "Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq." Nature **485**(7397): 201-206.

Doty, R. T., Vanasse, G. J., Disteche, C. M. and Willerford, D. M. (2002). "The leukemia-associated gene Mllt1/ENL: characterization of a murine homolog and demonstration of an essential role in embryonic development." Blood Cells Mol Dis **28**(3): 407-417.

Dray, E., Siaud, N., Dubois, E. and Doutriaux, M. P. (2006). "Interaction between Arabidopsis Brca2 and its partners Rad51, Dmc1, and Dss1." Plant Physiol **140**(3): 1059-1069.

Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T., et al. (2002). "Reversal of DNA alkylation damage by two human dioxygenases." Proc Natl Acad Sci U S A **99**(26): 16660-16665.

Edwards, C. A. and Ferguson-Smith, A. C. (2007). "Mechanisms regulating imprinted genes in clusters." Curr Opin Cell Biol **19**(3): 281-289.

Ekanayake, D. and Sabatini, R. (2011). "Epigenetic regulation of polymerase II transcription initiation in Trypanosoma cruzi: modulation of nucleosome abundance, histone modification, and polymerase occupancy by O-linked thymine DNA glucosylation." Eukaryot Cell **10**(11): 1465-1472.

Ekanayake, D. K., Minning, T., Weatherly, B., Gunasekera, K., Nilsson, D., et al. (2011). "Epigenetic regulation of transcription and virulence in Trypanosoma cruzi by O-linked thymine glucosylation of DNA." Mol Cell Biol **31**(8): 1690-1700.

Elkins, J. M., Hewitson, K. S., McNeill, L. A., Seibel, J. F., Schlemminger, I., et al. (2003). "Structure of factor-inhibiting hypoxia-inducible factor (HIF) reveals mechanism of oxidative modification of HIF-1 alpha." J Biol Chem **278**(3): 1802-1806.

Elkins, J. M., Ryle, M. J., Clifton, I. J., Dunning Hotopp, J. C., Lloyd, J. S., et al. (2002). "X-ray crystal structure of Escherichia coli taurine/alpha-ketoglutarate dioxygenase complexed to ferrous iron and substrates." Biochemistry **41**(16): 5185-5192.

Esteller, M. (2007). "Cancer epigenomics: DNA methylomes and histone-modification maps." Nat Rev Genet **8**(4): 286-298.

Falnes, P. O. (2004). "Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins." Nucleic Acids Res **32**(21): 6260-6267.

Falnes, P. O., Bjoras, M., Aas, P. A., Sundheim, O. and Seeberg, E. (2004). "Substrate specificities of bacterial and human AlkB proteins." Nucleic Acids Res **32**(11): 3456-3461.

Falnes, P. O., Johansen, R. F. and Seeberg, E. (2002). "AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli." Nature **419**(6903): 178-182.

Fernandez-Capetillo, O., Mahadevaiah, S. K., Celeste, A., Romanienko, P. J., Camerini-Otero, R. D., et al. (2003). "H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis." Dev Cell **4**(4): 497-508.

Ferrando, A. A., Armstrong, S. A., Neuberg, D. S., Sallan, S. E., Silverman, L. B., et al. (2003). "Gene expression signatures in MLL-rearranged T-lineage and B-precursor acute leukemias: dominance of HOX dysregulation." Blood **102**(1): 262-268.

Fischer, J., Koch, L., Emmerling, C., Vierkotten, J., Peters, T., et al. (2009). "Inactivation of the Fto gene protects from obesity." Nature **458**(7240): 894-898.

Fischle, W., Tseng, B. S., Dormann, H. L., Ueberheide, B. M., Garcia, B. A., et al. (2005). "Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation." Nature **438**(7071): 1116-1122.

Fodor, B. D., Kubicek, S., Yonezawa, M., O'Sullivan, R. J., Sengupta, R., et al. (2006). "Jmjd2b antagonizes H3K9 trimethylation at pericentric heterochromatin in mammalian cells." Genes Dev **20**(12): 1557-1562.

Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., et al. (2007). "A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity." Science **316**(5826): 889-894.

Frick, L. E., Delaney, J. C., Wong, C., Drennan, C. L. and Essigmann, J. M. (2007). "Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB." Proc Natl Acad Sci U S A **104**(3): 755-760.

Fu, D., Brophy, J. A., Chan, C. T., Atmore, K. A., Begley, U., et al. (2010a). "Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival." Mol Cell Biol **30**(10): 2449-2459.

Fu, Y., Dai, Q., Zhang, W., Ren, J., Pan, T., et al. (2010b). "The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA." Angew Chem Int Ed Engl **49**(47): 8885-8888.

Fujimoto, M., Izu, H., Seki, K., Fukuda, K., Nishida, T., et al. (2004). "HSF4 is required for normal cell growth and differentiation during mouse lens development." EMBO J **23**(21): 4297-4306.

Garcia-Cuellar, M. P., Zilles, O., Schreiner, S. A., Birke, M., Winkler, T. H., et al. (2001). "The ENL moiety of the childhood leukemia-associated MLL-ENL oncoprotein recruits human Polycomb 3." Oncogene **20**(4): 411-419.

Gerken, T., Girard, C. A., Tung, Y. C., Webby, C. J., Saudek, V., et al. (2007). "The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase." Science **318**(5855): 1469-1472.

Gould, A. (1997). "Functions of mammalian Polycomb group and trithorax group related genes." Curr Opin Genet Dev **7**(4): 488-494.

Grzyska, P. K., Muller, T. A., Campbell, M. G. and Hausinger, R. P. (2007). "Metal ligand substitution and evidence for quinone formation in taurine/alpha-ketoglutarate dioxygenase." J Inorg Biochem **101**(5): 797-808.

Gu, T. P., Guo, F., Yang, H., Wu, H. P., Xu, G. F., et al. (2011). "The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes." Nature **477**(7366): 606-610.

Gudmundsdottir, K., Lord, C. J., Witt, E., Tutt, A. N. and Ashworth, A. (2004). "DSS1 is required for RAD51 focus formation and genomic stability in mammalian cells." EMBO Rep **5**(10): 989-993.

Habu, T., Taki, T., West, A., Nishimune, Y. and Morita, T. (1996). "The mouse and human homologs of DMC1, the yeast meiosis-specific homologous recombination gene, have a common unique form of exon-skipped transcript in meiosis." Nucleic Acids Res **24**(3): 470-477.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., et al. (2002). "Epigenetic reprogramming in mouse primordial germ cells." Mech Dev **117**(1-2): 15-23.

Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., et al. (2010). "Crystal structure of the FTO protein reveals basis for its substrate specificity." Nature **464**(7292): 1205-1209.

Happel, N. and Doenecke, D. (2009). "Histone H1 and its isoforms: contribution to chromatin structure and function." Gene **431**(1-2): 1-12.

Hashimoto, H., Hong, S., Bhagwat, A. S., Zhang, X. and Cheng, X. (2012). "Excision of 5-hydroxymethyluracil and 5-carboxylcytosine by the thymine DNA glycosylase domain: its structural basis and implications for active DNA demethylation." Nucleic Acids Res.

Hashimoto, H., Liu, Y., Upadhyay, A. K., Chang, Y., Howerton, S. B., et al. (2012). "Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation." Nucleic Acids Res **40**(11): 4841-4849.

Hatfield, D. L., Carlson, B. A., Xu, X. M., Mix, H. and Gladyshev, V. N. (2006). "Selenocysteine incorporation machinery and the role of selenoproteins in development and health." Prog Nucleic Acid Res Mol Biol **81**: 97-142.

Hausinger, R. P. (2004). "FeII/alpha-ketoglutarate-dependent hydroxylases and related enzymes." Crit Rev Biochem Mol Biol **39**(1): 21-68.

He, C. (2010). "Grand challenge commentary: RNA epigenetics?" Nat Chem Biol **6**(12): 863-865.

He, N., Chan, C. K., Sobhian, B., Chou, S., Xue, Y., et al. (2011). "Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin." Proc Natl Acad Sci U S A **108**(36): E636-645.

He, Y. F., Li, B. Z., Li, Z., Liu, P., Wang, Y., et al. (2011). "Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA." Science **333**(6047): 1303-1307.

Hegg, E. L. and Que, L., Jr. (1997). "The 2-His-1-carboxylate facial triad--an emerging structural motif in mononuclear non-heme iron(II) enzymes." Eur J Biochem **250**(3): 625-629.

Hemenway, C. S., de Erkenez, A. C. and Gould, G. C. (2001). "The polycomb protein MPc3 interacts with AF9, an MLL fusion partner in t(9;11)(p22;q23) acute leukemias." Oncogene **20**(29): 3798-3805.

Henshaw, T. F., Feig, M. and Hausinger, R. P. (2004). "Aberrant activity of the DNA repair enzyme AlkB." J Inorg Biochem **98**(5): 856-861.

Hong, S., Cho, Y. W., Yu, L. R., Yu, H., Veenstra, T. D., et al. (2007). "Identification of JmjC domain-containing UTX and JMJD3 as histone H3 lysine 27 demethylases." Proc Natl Acad Sci U S A **104**(47): 18439-18444.

Hong, X., Zang, J., White, J., Wang, C., Pan, C. H., et al. (2010). "Interaction of JMJD6 with single-stranded RNA." Proc Natl Acad Sci U S A **107**(33): 14568-14572.

Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., et al. (2012). "PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse." Nucleic Acids Res **40**(Database issue): D261-270.

Hsia, D. A., Tepper, C. G., Pochampalli, M. R., Hsia, E. Y., Izumiya, C., et al. (2010). "KDM8, a H3K36me2 histone demethylase that acts in the cyclin A1 coding region to regulate cancer cell proliferation." Proc Natl Acad Sci U S A **107**(21): 9671-9676.

Hulse, J. D., Ellis, S. R. and Henderson, L. M. (1978). "Carnitine biosynthesis. beta-Hydroxylation of trimethyllysine by an alpha-ketoglutarate-dependent mitochondrial dioxygenase." J Biol Chem **253**(5): 1654-1659.

Hutton, J. J., Jr., Kaplan, A. and Udenfriend, S. (1967). "Conversion of the amino acid sequence gly-pro-pro in protein to gly-pro-hyp by collagen proline hydroxylase." Arch Biochem Biophys **121**(2): 384-391.

Hutton, J. J., Jr., Trappel, A. L. and Udenfriend, S. (1966). "Requirements for alpha-ketoglutarate, ferrous ion and ascorbate by collagen proline hydroxylase." Biochem Biophys Res Commun **24**(2): 179-184.

Iberg, A. N., Espejo, A., Cheng, D., Kim, D., Michaud-Levesque, J., et al. (2008). "Arginine methylation of the histone H3 tail impedes effector binding." J Biol Chem **283**(6): 3006-3010.

Ida, K., Kitabayashi, I., Taki, T., Taniwaki, M., Noro, K., et al. (1997). "Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13)." Blood **90**(12): 4699-4704.

Ido, A., Miura, Y. and Tamaoki, T. (1994). "Activation of ATBF1, a multiple-homeodomain zinc-finger gene, during neuronal differentiation of murine embryonal carcinoma cells." Dev Biol **163**(1): 184-187.

Iguchi-Ariga, S. M. and Schaffner, W. (1989). "CpG methylation of the cAMP-responsive enhancer/promoter sequence TGACGTCA abolishes specific factor binding as well as transcriptional activation." Genes Dev **3**(5): 612-619.

Iida, S., Seto, M., Yamamoto, K., Komatsu, H., Tojo, A., et al. (1993). "MLLT3 gene on 9p22 involved in t(9;11) leukemia encodes a serine/proline rich protein homologous to MLLT1 on 19p13." Oncogene **8**(11): 3085-3092.

Inoue, A., Shen, L., Dai, Q., He, C. and Zhang, Y. (2011). "Generation and replication-dependent dilution of 5fC and 5caC during mouse preimplantation development." Cell Res **21**(12): 1670-1676.

Inoue, A. and Zhang, Y. (2011). "Replication-dependent loss of 5-hydroxymethylcytosine in mouse preimplantation embryos." Science **334**(6053): 194.

Iqbal, K., Jin, S. G., Pfeifer, G. P. and Szabo, P. E. (2011). "Reprogramming of the paternal genome upon fertilization involves genome-wide oxidation of 5-methylcytosine." Proc Natl Acad Sci U S A **108**(9): 3642-3647.

Ito, S., D'Alessio, A. C., Taranova, O. V., Hong, K., Sowers, L. C., et al. (2010). "Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification." Nature **466**(7310): 1129-1133.

Ito, S., Shen, L., Dai, Q., Wu, S. C., Collins, L. B., et al. (2011). "Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine." Science **333**(6047): 1300-1303.

Ivan, M., Kondo, K., Yang, H., Kim, W., Valiando, J., et al. (2001). "HIFalpha targeted for VHL-mediated destruction by proline hydroxylation: implications for O2 sensing." Science **292**(5516): 464-468.

Jaakkola, P., Mole, D. R., Tian, Y. M., Wilson, M. I., Gielbert, J., et al. (2001). "Targeting of HIF-alpha to the von Hippel-Lindau ubiquitylation complex by O2-regulated prolyl hydroxylation." Science **292**(5516): 468-472.

Jia, G., Fu, Y., Zhao, X., Dai, Q., Zheng, G., et al. (2011). "N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO." Nat Chem Biol **7**(12): 885-887.

Jia, G., Yang, C. G., Yang, S., Jian, X., Yi, C., et al. (2008). "Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO." FEBS Lett **582**(23-24): 3313-3319.

Jones, P. A. (1999). "The DNA methylation paradox." Trends Genet **15**(1): 34-37.

Juszczynski, P., Rodig, S. J., Ouyang, J., O'Donnell, E., Takeyama, K., et al. (2010). "MLL-rearranged B lymphoblastic leukemias selectively express the immunoregulatory carbohydrate-binding protein galectin-1." Clin Cancer Res **16**(7): 2122-2130.

Kaelin, W. G., Jr. and Ratcliffe, P. J. (2008). "Oxygen sensing by metazoans: the central role of the HIF hydroxylase pathway." Mol Cell **30**(4): 393-402.

Kalhor, H. R. and Clarke, S. (2003). "Novel methyltransferase for modified uridine residues at the wobble position of tRNA." Mol Cell Biol **23**(24): 9283-9292.

Kataoka, H. and Sekiguchi, M. (1985). "Molecular cloning and characterization of the alkB gene of Escherichia coli." Mol Gen Genet **198**(2): 263-269.

Kataoka, H., Yamamoto, Y. and Sekiguchi, M. (1983). "A new gene (alkB) of Escherichia coli that controls sensitivity to methyl methane sulfonate." J Bacteriol **153**(3): 1301-1307.

Keeney, S., Baudat, F., Angeles, M., Zhou, Z. H., Copeland, N. G., et al. (1999). "A mouse homolog of the Saccharomyces cerevisiae meiotic recombination DNA transesterase Spo11p." Genomics **61**(2): 170-182.

Keeney, S., Giroux, C. N. and Kleckner, N. (1997). "Meiosis-specific DNA double-strand breaks are catalyzed by Spo11, a member of a widely conserved protein family." Cell **88**(3): 375-384.

Klose, R. J. and Bird, A. P. (2006). "Genomic DNA methylation: the mark and its mediators." Trends Biochem Sci **31**(2): 89-97.

Klose, R. J., Yamane, K., Bae, Y., Zhang, D., Erdjument-Bromage, H., et al. (2006). "The transcriptional repressor JHDM3A demethylates trimethyl histone H3 lysine 9 and lysine 36." Nature **442**(7100): 312-316.

Koivisto, P., Duncan, T., Lindahl, T. and Sedgwick, B. (2003). "Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase." J Biol Chem **278**(45): 44348-44354.

Koivisto, P., Robins, P., Lindahl, T. and Sedgwick, B. (2004). "Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases." J Biol Chem **279**(39): 40470-40474.

Kojic, M., Yang, H., Kostrub, C. F., Pavletich, N. P. and Holloman, W. K. (2003). "The BRCA2-interacting protein DSS1 is vital for DNA repair, recombination, and genome stability in Ustilago maydis." Mol Cell **12**(4): 1043-1049.

Kristensen, C. N., Bystol, K. M., Li, B., Serrano, L. and Brenneman, M. A. (2010). "Depletion of DSS1 protein disables homologous recombinational repair in human cells." Mutat Res **694**(1-2): 60-64.

Krivtsov, A. V., Feng, Z., Lemieux, M. E., Faber, J., Vempati, S., et al. (2008). "H3K79 methylation profiles define murine and human MLL-AF4 leukemias." Cancer Cell **14**(5): 355-368.

Kurowski, M. A., Bhagwat, A. S., Papaj, G. and Bujnicki, J. M. (2003). "Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB." BMC Genomics **4**(1): 48.

Laity, J. H., Lee, B. M. and Wright, P. E. (2001). "Zinc finger proteins: new insights into structural and functional diversity." Curr Opin Struct Biol **11**(1): 39-46.

Lan, F. and Shi, Y. (2009). "Epigenetic regulation: methylation of histone and non-histone proteins." Sci China C Life Sci **52**(4): 311-322.

Landini, P., Hajec, L. I. and Volkert, M. R. (1994). "Structure and transcriptional regulation of the Escherichia coli adaptive response gene aidB." J Bacteriol **176**(21): 6583-6589.
Landini, P. and Volkert, M. R. (2000). "Regulatory responses of the adaptive response to alkylation damage: a simple regulon with complex regulatory features." J Bacteriol **182**(23): 6543-6549.
Lando, D., Balmer, J., Laue, E. D. and Kouzarides, T. (2012). "The S. pombe histone H2A dioxygenase Ofd2 regulates gene expression during hypoxia." PLoS One **7**(1): e29765.
Lando, D., Peet, D. J., Gorman, J. J., Whelan, D. A., Whitelaw, M. L., et al. (2002). "FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor." Genes Dev **16**(12): 1466-1471.
Larsen, F., Gundersen, G., Lopez, R. and Prydz, H. (1992). "CpG islands as gene markers in the human genome." Genomics **13**(4): 1095-1107.
Laurent, L., Wong, E., Li, G., Huynh, T., Tsirigos, A., et al. (2010). "Dynamic changes in the human methylome during differentiation." Genome Res **20**(3): 320-331.
Lee, D. H., Jin, S. G., Cai, S., Chen, Y., Pfeifer, G. P., et al. (2005). "Repair of methylation damage in DNA and RNA by mammalian AlkB homologues." J Biol Chem **280**(47): 39448-39459.
Leonhardt, H., Page, A. W., Weier, H. U. and Bestor, T. H. (1992). "A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei." Cell **71**(5): 865-873.
Lin, C., Smith, E. R., Takahashi, H., Lai, K. C., Martin-Brown, S., et al. (2010). "AFF4, a component of the ELL/P-TEFb elongation complex and a shared subunit of MLL chimeras, can link transcription elongation to leukemia." Mol Cell **37**(3): 429-437.
Lindahl, T., Demple, B. and Robins, P. (1982). "Suicide inactivation of the E. coli O6-methylguanine-DNA methyltransferase." EMBO J **1**(11): 1359-1363.
Lister, R., Pelizzola, M., Dowen, R. H., Hawkins, R. D., Hon, G., et al. (2009). "Human DNA methylomes at base resolution show widespread epigenomic differences." Nature **462**(7271): 315-322.
Liu, A., Ho, R. Y., Que, L., Jr., Ryle, M. J., Phinney, B. S., et al. (2001). "Alternative reactivity of an alpha-ketoglutarate-dependent iron(II) oxygenase: enzyme self-hydroxylation." J Am Chem Soc **123**(21): 5126-5127.
Liu, W., Tanasa, B., Tyurina, O. V., Zhou, T. Y., Gassmann, R., et al. (2010). "PHF8 mediates histone H4 lysine 20 demethylation events involved in cell cycle progression." Nature **466**(7305): 508-512.
Loenarz, C. and Schofield, C. J. (2011). "Physiological and biochemical aspects of hydroxylations and demethylations catalyzed by human 2-oxoglutarate oxygenases." Trends Biochem Sci **36**(1): 7-18.
Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. and Richmond, T. J. (1997). "Crystal structure of the nucleosome core particle at 2.8 A resolution." Nature **389**(6648): 251-260.
Luger, K. and Richmond, T. J. (1998). "The histone tails of the nucleosome." Curr Opin Genet Dev **8**(2): 140-146.
Maiti, A. and Drohat, A. C. (2011). "Thymine DNA glycosylase can rapidly excise 5-formylcytosine and 5-carboxylcytosine: potential implications for active demethylation of CpG sites." J Biol Chem **286**(41): 35334-35338.
Martens, J. H., O'Sullivan, R. J., Braunschweig, U., Opravil, S., Radolf, M., et al. (2005). "The profile of repeat-associated histone lysine methylation states in the mouse epigenome." EMBO J **24**(4): 800-812.

Matthews, J. M. and Sunde, M. (2002). "Zinc fingers--folds for many occasions." IUBMB Life **54**(6): 351-355.
Mayer, W., Niveleau, A., Walter, J., Fundele, R. and Haaf, T. (2000). "Demethylation of the zygotic paternal genome." Nature **403**(6769): 501-502.
McCarthy, T. V., Karran, P. and Lindahl, T. (1984). "Inducible repair of O-alkylated DNA pyrimidines in Escherichia coli." EMBO J **3**(3): 545-550.
McDonough, M. A., Loenarz, C., Chowdhury, R., Clifton, I. J. and Schofield, C. J. (2010). "Structural studies on human 2-oxoglutarate dependent oxygenases." Curr Opin Struct Biol **20**(6): 659-672.
Meyer, C., Kowarz, E., Hofmann, J., Renneville, A., Zuna, J., et al. (2009). "New insights to the MLL recombinome of acute leukemias." Leukemia **23**(8): 1490-1499.
Miles, D. C., van den Bergen, J. A., Sinclair, A. H. and Western, P. S. (2010). "Regulation of the female mouse germ cell cycle during entry into meiosis." Cell Cycle **9**(2): 408-418.
Milne, T. A., Briggs, S. D., Brock, H. W., Martin, M. E., Gibbs, D., et al. (2002). "MLL targets SET domain methyltransferase activity to Hox gene promoters." Mol Cell **10**(5): 1107-1117.
Mishina, Y., Yang, C. G. and He, C. (2005). "Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins." J Am Chem Soc **127**(42): 14594-14595.
Mohan, M., Herz, H. M., Takahashi, Y. H., Lin, C., Lai, K. C., et al. (2010). "Linking H3K79 trimethylation to Wnt signaling through a novel Dot1-containing complex (DotCom)." Genes Dev **24**(6): 574-589.
Mohan, M., Lin, C., Guest, E. and Shilatifard, A. (2010). "Licensed to elongate: a molecular mechanism for MLL-based leukaemogenesis." Nat Rev Cancer **10**(10): 721-728.
Monsen, V. T., Sundheim, O., Aas, P. A., Westbye, M. P., Sousa, M. M., et al. (2010). "Divergent ss-hairpins determine double-strand versus single-strand substrate recognition of human AlkB-homologues 2 and 3." Nucleic Acids Res **38**(19): 6447-6455.
Muller, T. A., Meek, K. and Hausinger, R. P. (2010). "Human AlkB homologue 1 (ABH1) exhibits DNA lyase activity at abasic sites." DNA Repair (Amst) **9**(1): 58-65.
Namekawa, S. H., Park, P. J., Zhang, L. F., Shima, J. E., McCarrey, J. R., et al. (2006). "Postmeiotic sex chromatin in the male germline of mice." Curr Biol **16**(7): 660-667.
Neidig, M. L., Decker, A., Choroba, O. W., Huang, F., Kavana, M., et al. (2006). "Spectroscopic and electronic structure studies of aromatic electrophilic attack and hydrogen-atom abstraction by non-heme iron enzymes." Proc Natl Acad Sci U S A **103**(35): 12966-12973.
Nguyen, A. T., Taranova, O., He, J. and Zhang, Y. (2011). "DOT1L, the H3K79 methyltransferase, is required for MLL-AF9-mediated leukemogenesis." Blood **117**(25): 6912-6922.
Nordstrand, L. M., Furu, K., Paulsen, J., Rognes, T. and Klungland, A. (2012). "Alkbh1 and Tzfp repress a non-repeat piRNA cluster in pachytene spermatocytes." Nucleic Acids Res.
Nordstrand, L. M., Svard, J., Larsen, E., Nilsen, A., Ougland, R., et al. (2010). "Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects." PLoS One **5**(11): e13827.
Okano, M., Bell, D. W., Haber, D. A. and Li, E. (1999). "DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development." Cell **99**(3): 247-257.

Oswald, J., Engemann, S., Lane, N., Mayer, W., Olek, A., et al. (2000). "Active demethylation of the paternal genome in the mouse zygote." Curr Biol **10**(8): 475-478.
Ougland, R., Lando, D., Jonson, I., Dahl, J. A., Moen, M. N., et al. (2012). "ALKBH1 is a Histone H2A Dioxygenase Involved in Neural Differentiation." Stem Cells.
Ougland, R., Zhang, C. M., Liiv, A., Johansen, R. F., Seeberg, E., et al. (2004). "AlkB restores the biological function of mRNA and tRNA inactivated by chemical methylation." Mol Cell **16**(1): 107-116.
Pan, Z., Sikandar, S., Witherspoon, M., Dizon, D., Nguyen, T., et al. (2008). "Impaired placental trophoblast lineage differentiation in Alkbh1(-/-) mice." Dev Dyn **237**(2): 316-327.
Pastore, C., Topalidou, I., Forouhar, F., Yan, A. C., Levy, M., et al. (2012). "Crystal structure and RNA binding properties of the RNA recognition motif (RRM) and AlkB domains in human AlkB homolog 8 (ABH8), an enzyme catalyzing tRNA hypermodification." J Biol Chem **287**(3): 2130-2143.
Pfaffeneder, T., Hackner, B., Truss, M., Munzel, M., Muller, M., et al. (2011). "The discovery of 5-formylcytosine in embryonic stem cell DNA." Angew Chem Int Ed Engl **50**(31): 7008-7012.
Prendergast, G. C. and Ziff, E. B. (1991). "Methylation-sensitive sequence-specific DNA binding by the c-Myc basic region." Science **251**(4990): 186-189.
Qi, H. H., Sarkissian, M., Hu, G. Q., Wang, Z., Bhattacharjee, A., et al. (2010). "Histone H4K20/H3K9 demethylase PHF8 regulates zebrafish brain and craniofacial development." Nature **466**(7305): 503-507.
Ragvin, A., Valvatne, H., Erdal, S., Arskog, V., Tufteland, K. R., et al. (2004). "Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300." J Mol Biol **337**(4): 773-788.
Rando, O. J. (2012). "Combinatorial complexity in chromatin structure and function: revisiting the histone code." Curr Opin Genet Dev **22**(2): 148-155.
Reik, W. (2007). "Stability and flexibility of epigenetic gene regulation in mammalian development." Nature **447**(7143): 425-432.
Ringvoll, J., Moen, M. N., Nordstrand, L. M., Meira, L. B., Pang, B., et al. (2008). "AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA." Cancer Res **68**(11): 4142-4149.
Ringvoll, J., Nordstrand, L. M., Vagbo, C. B., Talstad, V., Reite, K., et al. (2006). "Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA." EMBO J **25**(10): 2189-2198.
Robinson, P. J. and Rhodes, D. (2006). "Structure of the '30 nm' chromatin fibre: a key role for the linker histone." Curr Opin Struct Biol **16**(3): 336-343.
Rubnitz, J. E., Morrissey, J., Savage, P. A. and Cleary, M. L. (1994). "ENL, the gene fused with HRX in t(11;19) leukemias, encodes a nuclear protein with transcriptional activation potential in lymphoid and myeloid cells." Blood **84**(6): 1747-1752.
Ryle, M. J., Liu, A., Muthukumaran, R. B., Ho, R. Y., Koehntop, K. D., et al. (2003). "O2- and alpha-ketoglutarate-dependent tyrosyl radical formation in TauD, an alpha-keto acid-dependent non-heme iron dioxygenase." Biochemistry **42**(7): 1854-1862.
Schueler, M. G. and Sullivan, B. A. (2006). "Structural and functional dynamics of human centromeric chromatin." Annu Rev Genomics Hum Genet **7**: 301-313.
Sedgwick, B., Bates, P. A., Paik, J., Jacobs, S. C. and Lindahl, T. (2007). "Repair of alkylated DNA: recent advances." DNA Repair (Amst) **6**(4): 429-442.
Sedgwick, B. and Lindahl, T. (2002). "Recent progress on the Ada response for inducible repair of DNA alkylation damage." Oncogene **21**(58): 8886-8894.

Semenza, G. L. (1999). "Regulation of mammalian O2 homeostasis by hypoxia-inducible factor 1." Annu Rev Cell Dev Biol **15**: 551-578.
Sharma, G. G., So, S., Gupta, A., Kumar, R., Cayrou, C., et al. (2010). "MOF and histone H4 acetylation at lysine 16 are critical for DNA damage response and double-strand break repair." Mol Cell Biol **30**(14): 3582-3595.
Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J. R., et al. (2004). "Histone demethylation mediated by the nuclear amine oxidase homolog LSD1." Cell **119**(7): 941-953.
Slany, R. K. (2009). "The molecular biology of mixed lineage leukemia." Haematologica **94**(7): 984-993.
Sleeman, M. C., Smith, P., Kellam, B., Chhabra, S. R., Bycroft, B. W., et al. (2004). "Biosynthesis of carbapenem antibiotics: new carbapenam substrates for carbapenem synthase (CarC)." Chembiochem **5**(6): 879-882.
Slotkin, R. K. and Martienssen, R. (2007). "Transposable elements and the epigenetic regulation of the genome." Nat Rev Genet **8**(4): 272-285.
Smith, E., Lin, C. and Shilatifard, A. (2011). "The super elongation complex (SEC) and MLL in development and disease." Genes Dev **25**(7): 661-672.
Songe-Moller, L., van den Born, E., Leihne, V., Vagbo, C. B., Kristoffersen, T., et al. (2010). "Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding." Mol Cell Biol **30**(7): 1814-1827.
Stender, J. D., Pascual, G., Liu, W., Kaikkonen, M. U., Do, K., et al. (2012). "Control of Proinflammatory Gene Programs by Regulated Trimethylation and Demethylation of Histone H4K20." Mol Cell.
Strahl, B. D. and Allis, C. D. (2000). "The language of covalent histone modifications." Nature **403**(6765): 41-45.
Sundheim, O., Talstad, V. A., Vagbo, C. B., Slupphaug, G. and Krokan, H. E. (2008). "AlkB demethylases flip out in different ways." DNA Repair (Amst) **7**(11): 1916-1923.
Sundheim, O., Vagbo, C. B., Bjoras, M., Sousa, M. M., Talstad, V., et al. (2006). "Human ABH3 structure and key residues for oxidative demethylation to reverse DNA/RNA damage." EMBO J **25**(14): 3389-3397.
Tahiliani, M., Koh, K. P., Shen, Y., Pastor, W. A., Bandukwala, H., et al. (2009). "Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1." Science **324**(5929): 930-935.
Tan, L. and Shi, Y. G. (2012). "Tet family proteins and 5-hydroxymethylcytosine in development and disease." Development **139**(11): 1895-1902.
Thalhammer, A., Bencokova, Z., Poole, R., Loenarz, C., Adam, J., et al. (2011). "Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1alpha (HIF-1alpha)." PLoS One **6**(1): e16210.
Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. and Sedgwick, B. (2002). "Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage." Nature **419**(6903): 174-178.
Tsujikawa, K., Koike, K., Kitae, K., Shinkawa, A., Arima, H., et al. (2007). "Expression and sub-cellular localization of human ABH family molecules." J Cell Mol Med **11**(5): 1105-1116.
Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M. E., Borchers, C. H., et al. (2006). "Histone demethylation by a family of JmjC domain-containing proteins." Nature **439**(7078): 811-816.

Tu, N., Hu, Y. and Mivechi, N. F. (2006). "Heat shock transcription factor (Hsf)-4b recruits Brg1 during the G1 phase of the cell cycle and regulates the expression of heat shock proteins." J Cell Biochem **98**(6): 1528-1542.

Tuck, M. T., Wiehl, P. E. and Pan, T. (1999). "Inhibition of 6-methyladenine formation decreases the translation efficiency of dihydrofolate reductase transcripts." Int J Biochem Cell Biol **31**(8): 837-851.

Turnbull, J. J., Nakajima, J., Welford, R. W., Yamazaki, M., Saito, K., et al. (2004). "Mechanistic studies on three 2-oxoglutarate-dependent oxygenases of flavonoid biosynthesis: anthocyanidin synthase, flavonol synthase, and flavanone 3beta-hydroxylase." J Biol Chem **279**(2): 1206-1216.

Vaillancourt, F. H., Yeh, E., Vosburg, D. A., O'Connor, S. E. and Walsh, C. T. (2005). "Cryptic chlorination by a non-haem iron enzyme during cyclopropyl amino acid biosynthesis." Nature **436**(7054): 1191-1194.

Vaillancourt, F. H., Yin, J. and Walsh, C. T. (2005). "SyrB2 in syringomycin E biosynthesis is a nonheme FeII alpha-ketoglutarate- and O2-dependent halogenase." Proc Natl Acad Sci U S A **102**(29): 10111-10116.

Vainio, S., Genest, P. A., ter Riet, B., van Luenen, H. and Borst, P. (2009). "Evidence that J-binding protein 2 is a thymidine hydroxylase catalyzing the first step in the biosynthesis of DNA base J." Mol Biochem Parasitol **164**(2): 157-161.

Valegard, K., van Scheltinga, A. C., Lloyd, M. D., Hara, T., Ramaswamy, S., et al. (1998). "Structure of a cephalosporin synthase." Nature **394**(6695): 805-809.

van den Born, E., Vagbo, C. B., Songe-Moller, L., Leihne, V., Lien, G. F., et al. (2011). "ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA." Nat Commun **2**: 172.

van Leeuwen, F., Taylor, M. C., Mondragon, A., Moreau, H., Gibson, W., et al. (1998). "beta-D-glucosyl-hydroxymethyluracil is a conserved DNA modification in kinetoplastid protozoans and is abundant in their telomeres." Proc Natl Acad Sci U S A **95**(5): 2366-2371.

Wang, G. L. and Semenza, G. L. (1995). "Purification and characterization of hypoxia-inducible factor 1." J Biol Chem **270**(3): 1230-1237.

Wang, Y., Wysocka, J., Sayegh, J., Lee, Y. H., Perlin, J. R., et al. (2004). "Human PAD4 regulates histone arginine methylation levels via demethylimination." Science **306**(5694): 279-283.

Webby, C. J., Wolf, A., Gromak, N., Dreger, M., Kramer, H., et al. (2009). "Jmjd6 catalyses lysyl-hydroxylation of U2AF65, a protein associated with RNA splicing." Science **325**(5936): 90-93.

Wei, Y. F., Carter, K. C., Wang, R. P. and Shell, B. K. (1996). "Molecular cloning and functional analysis of a human cDNA encoding an Escherichia coli AlkB homolog, a protein involved in DNA alkylation damage repair." Nucleic Acids Res **24**(5): 931-937.

Westbye, M. P., Feyzi, E., Aas, P. A., Vagbo, C. B., Talstad, V. A., et al. (2008). "Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA." J Biol Chem **283**(36): 25046-25056.

Whetstine, J. R., Nottke, A., Lan, F., Huarte, M., Smolikov, S., et al. (2006). "Reversal of histone lysine trimethylation by the JMJD2 family of histone demethylases." Cell **125**(3): 467-481.

Wossidlo, M., Nakamura, T., Lepikhov, K., Marques, C. J., Zakhartchenko, V., et al. (2011). "5-Hydroxymethylcytosine in the mammalian zygote is linked with epigenetic reprogramming." Nat Commun **2**: 241.

Xu, X. M. and Moller, S. G. (2011). "Iron-sulfur clusters: biogenesis, molecular mechanisms, and their functional significance." Antioxid Redox Signal **15**(1): 271-307.
Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., et al. (2006). "JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor." Cell **125**(3): 483-495.
Yang, C. G., Yi, C., Duguid, E. M., Sullivan, C. T., Jian, X., et al. (2008). "Crystal structures of DNA/RNA repair enzymes AlkB and ABH2 bound to dsDNA." Nature **452**(7190): 961-965.
Yao, T. P., Oh, S. P., Fuchs, M., Zhou, N. D., Ch'ng, L. E., et al. (1998). "Gene dosage-dependent embryonic development and proliferation defects in mice lacking the transcriptional integrator p300." Cell **93**(3): 361-372.
Yap, K. L. and Zhou, M. M. (2010). "Keeping it in the family: diverse histone recognition by conserved structural folds." Crit Rev Biochem Mol Biol **45**(6): 488-505.
Yi, C. and Pan, T. (2011). "Cellular dynamics of RNA modification." Acc Chem Res **44**(12): 1380-1388.
Yoder, J. A., Walsh, C. P. and Bestor, T. H. (1997). "Cytosine methylation and the ecology of intragenomic parasites." Trends Genet **13**(8): 335-340.
Yokoyama, A., Lin, M., Naresh, A., Kitabayashi, I. and Cleary, M. L. (2010). "A higher-order complex containing AF4 and ENL family proteins with P-TEFb facilitates oncogenic and physiologic MLL-dependent transcription." Cancer Cell **17**(2): 198-212.
Yoshida, K., Kondoh, G., Matsuda, Y., Habu, T., Nishimune, Y., et al. (1998). "The mouse RecA-like gene Dmc1 is required for homologous chromosome synapsis during meiosis." Mol Cell **1**(5): 707-718.
Youn, M. Y., Yokoyama, A., Fujiyama-Nakamura, S., Ohtake, F., Minehata, K., et al. (2012). "JMJD5, a Jumonji C (JmjC) domain-containing protein, negatively regulates osteoclastogenesis by facilitating NFATc1 protein degradation." J Biol Chem **287**(16): 12994-13004.
Yu, B., Edstrom, W. C., Benach, J., Hamuro, Y., Weber, P. C., et al. (2006). "Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB." Nature **439**(7078): 879-884.
Yu, Z., Genest, P. A., ter Riet, B., Sweeney, K., DiPaolo, C., et al. (2007). "The protein that binds to DNA base J in trypanosomatids has features of a thymidine hydroxylase." Nucleic Acids Res **35**(7): 2107-2115.
Zhang, Z., Ren, J., Stammers, D. K., Baldwin, J. E., Harlos, K., et al. (2000). "Structural origins of the selectivity of the trifunctional oxygenase clavaminic acid synthase." Nat Struct Biol **7**(2): 127-133.
Zheng, G., Dahl, J. A., Niu, Y., Fedorcsak, P., Huang, C. M., et al. (2012). "ALKBH5 Is a Mammalian RNA Demethylase that Impacts RNA Metabolism and Mouse Fertility." Mol Cell.
Zheng, L., Pan, H., Li, S., Flesken-Nikitin, A., Chen, P. L., et al. (2000). "Sequence-specific transcriptional corepressor function for BRCA1 through a novel zinc finger protein, ZBRK1." Mol Cell **6**(4): 757-768.
Zhu, J., He, F., Hu, S. and Yu, J. (2008). "On the nature of human housekeeping genes." Trends Genet **24**(10): 481-484.

I

BJ www.biochemj.org

Biochem. J. (2011) **434**, 391–398 (Printed in Great Britain) doi:10.1042/BJ20101667

Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4

Linn G. BJØRNSTAD, Giorgio ZOPPELLARO, Ane B. TOMTER, Pål Ø. FALNES[1] and K. Kristoffer ANDERSSON[1]

Department of Molecular Biosciences, University of Oslo, P.O. Box 1041, Blindern, NO-0316 Oslo, Norway

The Fe(II)/2OG (2-oxoglutarate)-dependent dioxygenase superfamily comprises proteins that couple substrate oxidation to decarboxylation of 2OG to succinate. A member of this class of mononuclear non-haem Fe proteins is the *Escherichia coli* DNA/RNA repair enzyme AlkB. In the present work, we describe the magnetic and optical properties of the yet uncharacterized human ALKBH4 (AlkB homologue). Through EPR and UV–visible spectroscopy studies, we address the Fe-binding environment of the proposed catalytic centre of wild-type ALKBH4 and an Fe(II)-binding mutant. We could observe a novel unusual Fe(III) high-spin EPR-active species in the presence of sulfide with a g_{max} of 8.2. The Fe(II) site was probed with NO. An intact histidine-carboxylate site is necessary for productive Fe binding. We also report the presence of a unique cysteine-rich motif conserved in the N-terminus of ALKBH4 orthologues, and investigate its possible Fe-binding ability. Furthermore, we show that recombinant ALKBH4 mediates decarboxylation of 2OG in absence of primary substrate. This activity is dependent on Fe as well as on residues predicted to be involved in Fe(II) co-ordination. The present results demonstrate that ALKBH4 represents an active Fe(II)/2OG-dependent decarboxylase and suggest that the cysteine cluster is involved in processes other than Fe co-ordination.

Key words: AlkB, AlkB homologue (ALKBH4), EPR, non-haem Fe, UV–visible spectroscopy.

INTRODUCTION

The superfamily of Fe(II) and 2OG (2-oxoglutarate, also known as α-ketoglutarate)-dependent dioxygenases (Pfam accession number PF03171) is the largest known non-haem Fe protein family able to carry out hydroxylation reactions of unactivated C–H groups. These enzymes act on a variety of substrates [1,2], and the reaction occurs by reductive activation of molecular oxygen coupled with decarboxylation of the co-substrate 2OG to succinate [3–5]. Within the oxidation process through an Fe(IV)O intermediate, one of the oxygen atoms from O_2 is incorporated into the succinate moiety and the other becomes a hydroxy group in the product [3–5]. The catalytically active site is formed by a mononuclear non-haem Fe centre co-ordinated, in general, by two histidine residues and one carboxylate moiety [6–9]. This site is responsible for the binding of 2OG as well as dioxygen. During turnover, the Fe(II) hexa-co-ordinated states change to a final penta-co-ordinated state, with an open oxygen co-ordination site [6–9], as observed in the three-dimensional structure of the deacetoxycephalosporin C synthase with Fe(II) and 2OG bound (PDB code 1RXG), the first structurally characterized member of this family [10]. However, when uncoupled turnover of 2OG takes place, either in the absence of natural substrates or due to incorrect orientation of substrates in the active site, decomposition of 2OG into succinate and CO_2 may lead to enzyme deactivation.

The Fe(II)/2OG dioxygenase AlkB from *Escherichia coli* is a repair enzyme, which is induced as part of the adaptive response to alkylation damage. AlkB catalyses demethylation of 1meA (1-methyladenine) and 3meC (3-methylcytosine) in DNA in a reaction where the methyl group is hydroxylated and then released as formaldehyde, thereby regenerating the normal base [11,12]. Additionally, AlkB repairs the structurally analogous lesions 1meG (1-methylguanine) and 3meT (3-methylthymidine) and the bulkier exocyclic etheno and ethano adducts [13–17]. Moreover, AlkB also displays activity on methylated RNA [18]. A bioinformatics analysis revealed eight different mammalian ALKBHs (AlkB homologues), denoted ALKBH1–8 [19], and a more recent study has demonstrated FTO (fat mass and obesity-associated protein) [20] to be a functional ALKBH and thus the ninth member of this family [21]. AlkB-like *in vitro* repair activities have been reported for ALKBH1 [22], ALKBH2, ALKBH3 [23] and FTO [21], but only ALKBH2 has been convincingly demonstrated to function as a repair enzyme *in vivo* [24,25].

Recently, it has become clear that the ALKBH proteins are involved in processes other than DNA/RNA repair, as ALKBH8 has been demonstrated to be a tRNA modification enzyme [26–28] and ALKBH1 has been implicated in gene regulation [29]. ALKBH4–ALKBH7 remain completely uncharacterized; in fact, nothing is known about their ability to bind Fe-metal ion, the nature of the so-formed active site(s) and their optical/magnetic fingerprints. From the sequence alignment of putative ALKBH4 orthologues from various organisms, we noted the presence of a cluster of four highly conserved cysteine residues very close to the N-terminus (Figure 1). This cluster is not present in AlkB, and we reasoned that it potentially constitutes a complementary Fe–S binding site. In order to clearly elucidate the nature of the Fe-metal binding by human ALKBH4 in its anticipated binding triad of two histidine residues and one carboxylate motif (His^{169}-Asp^{171}-His^{254}) (Figure 2) and to address the possible role of the cysteine cluster as an additional Fe–S centre, in the present work we studied the enzymatic activities and the optical and magnetic fingerprints

Abbreviations used: ALKBH, AlkB homologue; FTO, fat mass and obesity-associated protein; GST, glutathione transferase; ICP-AEP, inductively coupled plasma atomic emission spectroscopy; IPNS, isopenicillin N synthase; IPTG, isopropyl β-D-thiogalactopyranoside; $MV^{\bullet+}$, Methyl Viologen radical cation; 2OG, 2-oxoglutarate; PAH, phenylalanine hydroxylase; 4,5-PCD, protocatechuate 4,5-dioxygenase; TauD, taurine dioxygenase; UV–Vis, UV–visible; ZFS, zero-field splitting.

[1] Correspondence may be addressed to either of these authors (email k.k.andersson@imbv.uio.no or pal.falnes@imbv.uio.no).

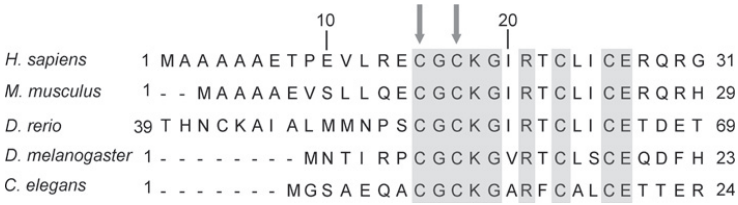


Figure 1 Sequence alignment of the N-terminal part of putative ALKBH4 orthologues from different organisms

Arrows indicate the cysteine residues subjected to mutagenesis (C15A/C17A). GenBank® Identifier (gi) numbers of the displayed sequences are as follows: 8923019 (*Homo sapiens*), 81905645 (*Mus musculus*), 68372246 (*Danio rerio*), 24583140 (*Drosophila melanogaster*) and 25148697 (*Caenorhabditis elegans*). The alignment was generated using the MAFFT algorithm [48] and displayed in Jalview [49].

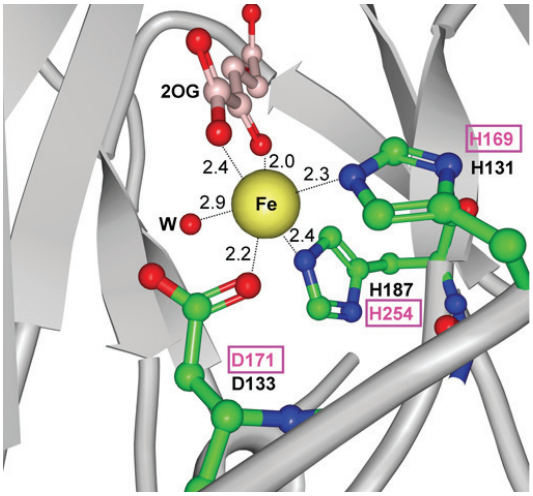


Figure 2 Close-up view of the catalytic centre of AlkB in the presence of bound 2OG (PDB code 2FD8) [50]

The amino acid residues that complex the Fe metal are indicated in black (H131, D133, H187). A water molecule (W) weakly interacting with the Fe ion (Fe–H_2O = 2.9 Å; 1 Å = 0.1 nm) is also shown. The amino acid residues of ALKBH4 thought to form a similar Fe metal-binding triad (H169, D171, H254) [19] are indicated by boxed magenta labels.

[UV–Vis (UV–visible)/EPR] of the wild-type protein and of two mutant derivatives. One mutant (C15A/C17A) contained an abrogated cysteine cluster, whereas the motif putatively involved in Fe(II) binding was destroyed in the other (H169A/D171A).

MATERIALS AND METHODS

Plasmid construction and site-specific mutagenesis of ALKBH4

The cDNA encoding human ALKBH4 was amplified by PCR and inserted between the NdeI and BamHI sites of the expression vector pET-28a (Novagen). The QuikChange™ Site-Directed Mutagenesis Kit (Stratagene) was used for generation of ALKBH4 mutants. To generate the pET-H169A/D171A plasmid, the primers 5′-CTGCCATTGACCCCGCCCTGGCCGACGCCTGGCTG-3′ and 5′-CAGCCAGGCGTCGGCCAGGGCGGGGTCAATGGCAG-3′ were used, and the primers 5′-CGAAGTCCTTCGGGAAGCCGGTGCCAAGGGCATCCGGACC-3′ and 5′-GGTCCGGATGCCCTTGGCACCGGCTTCCCGAAGGACTTCG-3′ were used for generation of the plasmid pET-C15A/C17A (underlining indicates mutation sites). Plasmids encoding GST (glutathione transferase)-tagged mutants were obtained by subcloning the mutated cDNA fragments into the vector pGEX-6P-2 (GE Healthcare), resulting in the plasmids pGEX-H169A/D171A and pGEX-C15A/C17A respectively (BamHI and EcoRI sites were used for H169A/D171A, and BamHI and SalI sites were used for C15A/C17A). All PCR-generated fragments were verified by DNA sequencing.

Purification of ALKBH4 and mutant proteins

Plasmids were transformed into the *E. coli* strain BL21-CodonPlus(DE3)-RIPL (Stratagene), and protein expression was induced by the addition of 0.075–0.1 mM IPTG (isopropyl β-D-thiogalactopyranoside). Cells were disrupted by French press treatment [two passages through a standard FRENCH® pressure cell (Thermo Electron) at a working pressure of 1000 psi (1 psi = 6.9 kPa)], and proteins were purified from clarified lysates using affinity purification. For spectroscopic analysis, ALKBH4 and mutants were expressed as GST-fusion proteins and purified using glutathione–Sepharose 4B medium (GE Healthcare). The GST tag was removed by on-column enzymatic cleavage using PreScission protease (GE Healthcare), and HiTrap Desalting columns (GE Healthcare) were used for buffer exchange (50 mM Hepes, pH 7.0, and 100 mM KCl). His_6-tagged AlkB proteins for decarboxylation assays were purified using TALON® Metal Affinity Resin (Clontech). Proteins were analysed by SDS/4–12 % PAGE and subsequent Coomassie Brilliant Blue staining. The metal content (Fe) in the protein preparations was probed by ICP-AEP (inductively coupled plasma atomic emission spectroscopy), with detection at 238.204 nm for the Fe ion emission line.

Assay for decarboxylation of 2OG to succinate

The reactions were performed as described previously [14]. Briefly, reaction mixtures of 50 μl contained 50 mM Hepes, pH 7.5, 4 mM ascorbic acid, 160 μM 2OG (10 % [5–^{14}C]2OG), 80 μM $FeSO_4$ (when indicated) and His_6-tagged ALKBH4 (500 pmol of wild-type or mutants) or *E. coli* His–AlkB (100 pmol). EDTA (1 mM) was added to samples without $FeSO_4$. Reactions were incubated at 37 °C for 30 min. The amount of [1–^{14}C]succinate formed was measured by scintillation counting after precipitation of the remaining 2OG with 2,4-DNPH (2,4-dinitrophenylhydrazine).

Metal-reconstitution procedures and spectroscopic measurements

Concentrated (<300 μM) protein solutions were obtained by centrifugal filter devices (Microcon YM-10) in 50 mM Hepes, pH 7.0, and were then made anaerobic in airtight vessels by several rounds of vacuum treatment and argon exchange using the Schlenk technique [29a]. The minimum amount ($\leqslant$2 mol per mol of protein) of a freshly prepared anaerobic solution of $Na_2S_2O_4$ (5.7 mM in 50 mM Hepes, pH 7.0) was then added to the protein solutions using airtight syringes. Stoichiometric amounts of Fe(II) ammonium sulfate were added to the protein samples in a similar way, using anaerobic stocks. In some cases, small amounts ($\sim$20 μM) of Methyl Viologen were added in order to monitor that strict anaerobic conditions were maintained within reconstitution procedures. Anaerobic solutions of 2OG and Na_2S in 50 mM Hepes, pH 7.0, were freshly prepared and added to the protein samples at a 10-fold excess. These solutions were transferred to quartz EPR tubes through airtight syringes that were maintained under argon, and then the tubes were sealed with Precision Seal® rubber septa (Sigma–Aldrich). Nitric oxide (NO) gas was added through airtight syringes directly into the quartz EPR tubes containing the sample protein, or to airtight 1 cm optical path quartz cuvettes. Samples were maintained at cold temperatures ($\leqslant$10 °C). Sample oxidation was obtained by air exposure followed by slow injection of O_2 ($\sim$90 % purity, $\sim$5 min). In the experiments employing reconstitution in the presence of 2OG/Na_2S/$Na_2S_2O_4$, sample oxidation was obtained by injection of 50 μl of O_2-saturated 50 mM Hepes, pH 7.0, into EPR tubes containing 100 μl of protein sample, followed by fast quenching in liquid nitrogen. The low-temperature EPR spectra were acquired on a Bruker Elexsys 560 instrument equipped with a dual-band X-resonator (Bruker) and by using a ESR900 He-flow cryostat (Oxford Instruments). UV–Vis spectra were recorded on a HP 8452A diode-array spectrophotometer (Hewlett Packard).

RESULTS AND DISCUSSION

Decarboxylation activity of ALKBH4 and mutant proteins towards 2OG

Like several of the Fe(II)/2OG dioxygenases, *E. coli* AlkB possesses the ability to perform decarboxylation of 2OG to succinate even in the absence of primary substrate [11,12]. Since the primary substrate of ALKBH4 is still unknown, we took advantage of such uncoupled co-substrate conversion to test the *in vitro* enzymatic activity of ALKBH4. For this purpose, recombinant His_6-tagged ALKBH4 and two mutants were purified from *E. coli* (Figure 3A) and subsequently assayed for 2OG decarboxylation activity. Indeed, ALKBH4 was able to catalyse succinate formation in the presence of the co-factor Fe(II), although not to the same extent as AlkB (Figure 3B). The decarboxylation activity was reduced to background levels when Fe(II) was not included, demonstrating the requirement for Fe, which is indicative of AlkB proteins. Likewise, the H169A/D171A mutant was devoid of the ability to perform 2OG turnover. The C15A/C17A mutant displayed a decarboxylation activity similar to that of the wild-type ALKBH4 protein, suggesting that the cysteine motif is not involved in Fe(II) co-ordination. Taken together, these results experimentally confirmed that ALKBH4 is a Fe(II)/2OG-dependent decarboxylase.

Spectroscopic fingerprints of the Fe catalytic centre of ALKBH4 and mutant proteins probed by EPR

The electronic/magnetic fingerprints of the native and mutant proteins were then addressed by low-temperature EPR

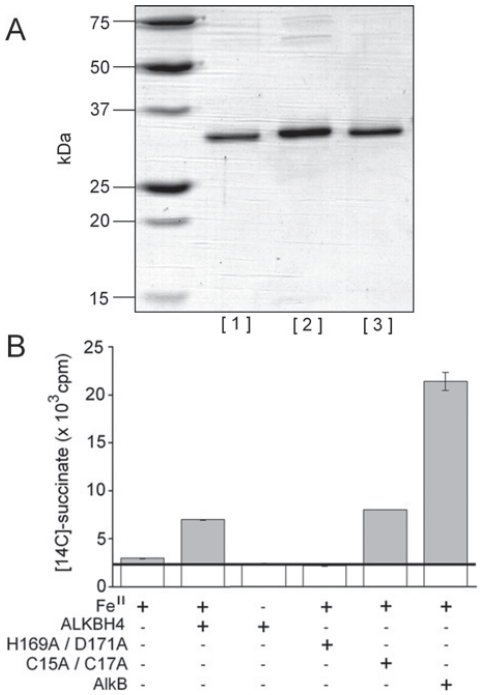


Figure 3 2OG decarboxylation activity of ALKBH4

(**A**) SDS/PAGE (4–12 %) analysis of purified recombinant proteins. Lane 1, ALKBH4; lane 2, H169A/D171A mutant; and lane 3, C15A/C17A mutant. The molecular mass in kDa is indicated on the left-hand side. (**B**) ALKBH4 (500 pmol of wild-type or mutant protein) or AlkB (100 pmol) was incubated in the presence of [5–^{14}C]2OG and the resulting [^{14}C]succinate formed was measured by scintillation counting. Fe(II) was added as indicated. Results are means $\pm$ S.D. of duplicate samples. The bold line shows the averaged background value (2500 c.p.m.).

measurements down to cryogenic temperatures. We probed different conditions where the metal co-ordination environment of Fe reconstituted ALKBH4, and mutant proteins were analysed both in the reduced [Fe(II)] and oxidized [Fe(III)] states, as well as in the presence or absence of 2OG/succinate. The results are summarized in Table 1. It should be noted that the isolated proteins (ALKBH4 and mutants), even when kept in iced-cooled solutions (4 °C), aggregated and then precipitated at concentrations higher than 0.5 mM. The process was accelerated greatly when an excess (e.g. 5-fold) of dithionite was added to the solution when anaerobic Fe(II)-reconstitution procedures were carried out. Nevertheless, low protein concentrations (<0.3 mM) and a slight excess of $Na_2S_2O_4$ (approx. twice greater) allowed efficient reconstitution of the apo-proteins with the Fe(II) metal ion, hence avoiding precipitation. The purified ALKBH4 protein, in the absence of added cofactors, gave rise to an extremely weak high-spin Fe(III) signal at $g \approx 4.3$ (Figure 4A). This resonance did not increase upon addition of oxidants (e.g. Na_2IrCl_6). This agrees with the ICP-AES findings that indicated the presence of Fe traces only in the purified protein. Traces of Cu(II) ion as an impurity (around $g \approx 2.1$) were also detected, with relative amounts that changed from sample to sample. When ALKBH4 and the C15A/C17A and H169A/D171A mutants were incubated with stoichiometric amounts of Fe(II) under reducing conditions, either in the presence or absence of 2OG/succinate, the recorded

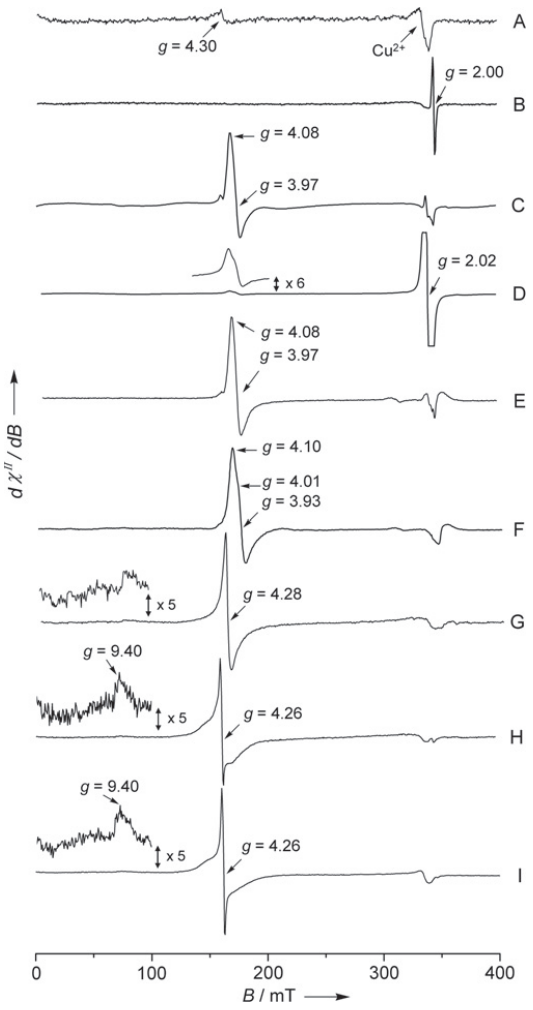


Figure 4 EPR spectroscopy of Fe-reconstituted ALKBH4 and mutant proteins

X-band EPR spectra of (**A**) purified ALKBH4, (**B**) ALKBH4 after incubation with stoichiometric amounts of Fe(II), (**C**) ALKBH4/Fe(II) after addition of NO, (**D**) the H169A/D171A/Fe(II) mutant after addition of NO, (**E**) ALKBH4/Fe(II) after addition of NO in the presence of 2OG, (**F**) ALKBH4/Fe(II) after addition of NO in the presence of succinate, (**G**) the H169A/D171A/Fe(II) mutant after addition of O_2, (**H**) ALKBH4/Fe(II)/2OG after addition of O_2 and (**I**) ALKBH4/Fe(II) after addition of O_2. Protein concentrations were 180–200 μM in 50 mM Hepes, pH 7.0. The measurements were performed at $T = 8.0 \pm 0.5$ K (approx. −265 °C) with the following parameters: frequency, 9.67 GHz; modulation frequency, 100 KHz; modulation amplitude, 0.7 mT; gain, 55 dB; time constant, 81.92 ms; sweep time, 323 s; microwave power, 0.4–0.8 mW. Two scans were accumulated and averaged; cavity quality factor $Q = 4300$–4600. Note that in (**B**) the sharp signal at $g = 2.00$ arises from $MV^{\bullet+}$, which was used as an internal indicator to probe the attainment of rigorous anaerobic conditions.

low-temperature EPR spectra did not show any Fe-related signal (Figure 4B). It is important to note that the only resonance signal emerging in the spectra was always associated with the presence of $MV^{\bullet+}$ (Methyl Viologen radical cation) ($g = 2.00$ in Figure 4B), which was used as an internal indicator to probe the attainment of rigorous anaerobic conditions. These results can be interpreted at X-band EPR with either (i) an integer-spin system having a high-spin Fe(II) state ($S = 2$, where S is the ground-state spin) characterized by a large axial ZFS (zero-field splitting) term called δ in EPR or Δ in MCD (magnetic CD), as it occurs in many mononuclear non-haem Fe proteins where such configuration is usually not spectroscopically accessible, or alternatively (ii) by a low-spin Fe(II) state ($S = 0$) [1,2,30]. NO has been successfully used as a probe for such centres [31], since it can convert the $S = 2$ species into an EPR-active $S = 3/2$ species, as observed, for example, in IPNS (isopenicillin N synthase) [32], in extradiol-cleaving catechol dioxygenases such as 2,3-CTD (catechol 2,3-dioxygenase) and 4,5-PCD (protocatechuate 4,5-dioxygenase) [33,34], as well as in non-haem Fe model complexes [35]. Upon anaerobic addition of NO to Fe(II)-reconstituted ALKBH4 and C15A/C17A protein in the absence of 2OG, a strong axial EPR signal centred at a g_{eff} of 3.97 developed (g-tensor components at $g_1 = 4.08$, $g_2 = 3.97$ and $g_3 = 2.00$), together with a second asymmetric resonance at $g \approx 2.00$ (Figure 4C). The latter signal contained features characteristic of a small amount of NO radical being trapped in the frozen matrix solution. Similarly, when anaerobic reconstitution of the ALKBH4 and C15A/C17A proteins was carried out with both Fe(II) and a 10-fold excess of 2OG, a strong axial resonance signal centred at $g_{eff} = 3.97$ developed after NO addition, being accompanied by a small highly asymmetric signal around $g \approx 2.00$ (Figure 4E). Analogous spurious signals at $g \approx 2.00$ are found in other non-haem Fe proteins, such as 4,5-PCD [36] and IPNS [32]. Therefore in both cases a nearly axial $S = 3/2$ Fe(II)–NO adduct formed with the E (rhombic) versus D (axial) term $|E/D| \approx 0.010$.

Similarly, anaerobic reconstitution of the proteins with both Fe(II) and a 10-fold excess of the product succinate, followed by addition of NO, led to a single strong resonance signal around $g_{eff} = 3.93$ that contained a slightly higher degree of rhombicity ($|E/D| \approx 0.015$) (Figure 4F). The H169A/D171A protein behaved very differently. After being incubated with Fe(II) under anaerobic conditions, a very strong signal centred at $g \approx 2$ appeared immediately after addition of NO. Only a very weak resonance around $g \approx 4$ could be detected, with small variations in relative intensities from preparation to preparation. These findings confirm that when the histidine-carboxylate Fe-binding region is abrogated, only a much smaller fraction of Fe productively forms stable $S = 3/2$ Fe–NO complexes (Figure 4D). Upon addition of oxygen, the ALKBH4 protein sample that contained 2OG (Figure 4H), and also the sample that was deficient in co-substrate (Figure 4I), displayed formation of strong similar signals with broad shoulders at $g_{eff} = 4.26$, being accompanied by the appearance of weak resonances at $g_{eff} = 9.40$. Those signals originate from transition involving the middle and ground Kramer's doublets respectively and are typically found in high-spin rhombic Fe(III) centres ($|E/D| \approx 0.25$). The same resonances developed in the C15A/C17A mutant upon Fe(II) oxidation. Similarly, a strong signal formed at $g_{eff} = 4.28$ upon oxidation of the Fe(II) H169A/D171A mutant (Figure 4G). However, the presence of broad shoulders like those of the Fe(III) forms of both ALKBH4 and the C15A/C17A mutant were not detected. Such EPR resonance can therefore be interpreted as arising from spurious or non-specifically bound Fe(III) metal ion. Surprisingly, addition of oxygen to the metal-reconstituted ALKBH4 and C15A/C17A proteins in the presence of succinate resulted in protein precipitation. The reason for this is presently unknown.

Motivated by the possibility that the N-terminal cysteine-rich motif might provide an Fe–S cluster, we measured the EPR spectrum of ALKBH4 in the presence of inorganic S and Fe(II).

Table 1 X-band EPR data obtained at $T = 8.0 \pm 0.5$ K for ALKBH4 and mutant proteins (H169A/D171A and C15A/C17A), in their purified forms, after Fe-metal reconstitution in the presence or absence of 2OG/succinate and upon interaction with NO

D, axial ZFS term; *E*, rhombic ZFS term; n.d., not determined; *S*, electronic spin state associated with the individual paramagnetic centre (in parentheses/square brackets).

Protein	Observed EPR *g*-values	$\|E/D\|$	*S*	Notes
ALKBH4 purified[a,b]	4.30 [Fe(III)], 2.10 [Cu(II)]	~1/3	5/2 [Fe(III)], 1/2 [Cu(II)]	Fe(III) and Cu(II) traces
ALKBH4/Fe(II)[a,b]	2.00	n.d.	2 [Fe(II)], 1/2 ($MV^{\bullet+}$)	Recorded under reducing conditions ($Na_2S_2O_4$) in the presence of $MV^{\bullet+}$. Fe(II) was not EPR-visible at X-band; $MV^{\bullet+}$ was EPR active
ALKBH4/Fe(II)/NO[a]	4.08, 3.97, 2.00	0.010	3/2 [Fe(II)–NO complex]	Recorded under reducing conditions ($Na_2S_2O_4$)
ALKBH4/Fe(II)/2OG/NO[a]	4.08, 3.97, 2.00	0.010	3/2 [Fe(II)–NO complex]	Recorded under reducing conditions ($Na_2S_2O_4$)
ALKBH4/Fe(II)/succinate/NO[a]	4.10, 3.93, 2.00	0.015	3/2 [Fe(II)–NO complex]	Recorded under reducing conditions ($Na_2S_2O_4$)
H169A/D171A/Fe(II)/NO	~4.00, 2.02	n.d.	3/2 [Fe(II)–NO complex], 1/2 ($NO^{\bullet}$ radical)	Recorded under reducing conditions ($Na_2S_2O_4$); $NO^{\bullet}$ radical was the major species, Fe(II)–NO complex was a minor species
ALKBH4/Fe(III)/2OG/Na_2S[a]	8.24, 5.56, 4.26, 2.10	0.120	5/2 [Fe(III)], 1/2 [Cu(II)]	Transient species within the Fe(II)→Fe(III) oxidation process. Cu(II) was a minor species
ALKBH4/Fe(III)[a]	9.40, 4.26, 2.10	0.250	5/2 [Fe(III)], 1/2 [Cu(II)]	Cu(II) was a minor species
H169A/D171A/Fe(III)	4.28, 2.10	1/3	5/2 [Fe(III)], 1/2 [Cu(II)]	Non-specifically bound Fe(III). Cu(II) was a minor species
ALKBH4/Fe(III)/2OG[a]	9.40, 4.26, 2.10	0.250	5/2 [Fe(III)], 1/2 [Cu(II)]	Cu(II) was a minor species
ALKBH4/Fe(III)/succinate[a]	n.d.	n.d.	n.d.	Protein precipitation

[a]The same features were observed for the C15A/C17A mutant protein.
[b]The same features were observed for the H169A/D171A mutant protein.

Anaerobic incubation of ALKBH4 with 2OG and Fe(II) in the presence of Na_2S (10-fold excess), followed by addition of O_2 and fast quenching at liquid-N_2 temperature (77 K, –196 °C), resulted in a transient EPR-active species that exhibited resonances at $g_{eff} = 8.24$, 5.56 and 4.26 (Figure 5). Upon thawing, this new species relaxed fast (~20 s) to the high-spin rhombic Fe(III) ion with $|E/D| \approx 0.25$, showing resonances similar to those described in Figure 4(H). Identical behaviour was observed for the C15A/C17A mutant, but not for the mutant H169A/D171A. These effects suggest that the structural changes occuring in the active site within the Fe(II)→Fe(III) oxidation process are modulated further by the presence of Na_2S. It is important to note that similar signals and a similar decay process have been reported in the non-haem Fe(III) centre present in photosystem II, for example, decrease of the $g_{eff} = 8.07$ and $g_{eff} = 5.58$ signals being accompanied by induction (increase) of the signal at the $g_{eff} = 4.26$ signal, as a result of UV-B irradiation (280–320 nm) [37,38]. These types of resonances have previously been analysed by Weisser et al. [39] in terms of ZFS-distributed $S = 5/2$ systems in Fe–catecholato complexes containing moderate rhombic distortion ($|E/D| \approx 0.12$). Furthermore, EPR spectra similar to those reported in Figure 5, although not transient, are observed in Fe(III)–catecholato complexes as model systems of intradiol-cleaving catechol dioxygenases [40], in PAH (phenylalanine hydroxylase) in the presence of catecholamine feedback inhibitors or even when recording PAH in Tris buffer [41,42].

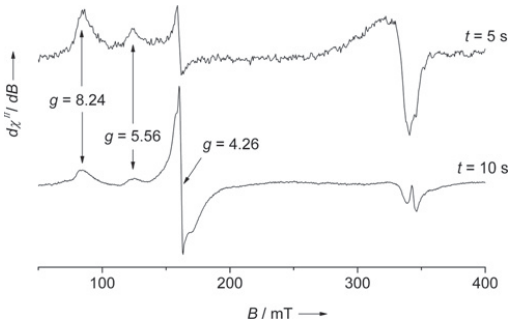


Figure 5 Transient EPR spectra of ALKBH4 recorded within the Fe(II)→Fe(III) oxidation process

X-band EPR spectra of ALKBH4 (120 μM in 50 mM Hepes, pH 7.0) reconstituted with stoichiometric amounts of Fe(II), 2OG (10-fold excess) and Na_2S (10-fold excess), recorded after addition of O_2 followed by quenching after $t = 5$ s (upper trace) or $t = 10$ s (lower trace) of the ice-cooled solutions down to liquid-N_2 temperature ($T = 77$ K). The measurements were performed at $T = 8.0 \pm 0.5$ K; frequency, 9.67 GHz; modulation frequency, 100 KHz; modulation amplitude, 0.7 mT; gain, 55 dB; time constant, 81.92 ms; sweep time, 323 s; microwave power, 0.8 mW; cavity quality factor $Q = 4200$–4400; two scans were accumulated and averaged.

The optical signatures of ALKBH4 and mutant proteins probed by UV–Vis absorption spectroscopy

In parallel with the low-temperature EPR experiments, the optical fingerprints of the various forms of ALKBH4 and mutants [Fe(II)/Fe(III) in the presence of, and without, 2OG] were probed by UV–Vis absorption spectroscopy. The electronic spectrum of the purified ALKBH4 protein did not show absorption features in the visible region (Figure 6A, dashed line). Similarly, the C15A/C17A and H169A/D171A proteins gave featureless spectra. Upon reconstitution with Fe(II), but in the absence of 2OG, both ALKBH4 and the C15A/C17A mutant developed a broad and composite absorption band, extending from 400 nm ($\lambda = 404$ nm, $\varepsilon \approx 570$ $M^{-1} \cdot cm^{-1}$) to 600 nm ($\lambda = 490$ nm, $\varepsilon \approx 260$ $M^{-1} \cdot cm^{-1}$) (Figure 6A, dashed and dotted line). However, the strong absorption band around 320 nm, due to the presence of $Na_2S_2O_4$ (reducing equivalents) in the medium, masked an additional high-energy absorption signature, around 350 nm ($\varepsilon \approx 920$ $M^{-1} \cdot cm^{-1}$), as substantiated in those samples reconstituted with Fe(II) without the supply of reducing agents (Figure 6A, continuous line). When a 10-fold excess of 2OG was added under reducing conditions, a slightly different chromophore ($\lambda = 430$ nm, $\varepsilon \approx 440$ $M^{-1} \cdot cm^{-1}$) developed quickly in both ALKBH4/Fe(II) and the C15A/C17A/Fe(II) protein (Figure 6B, dotted line), with an absorption envelope similar to that observed by Henshaw et al. [43] in *E. coli* AlkB in the presence of Fe(II)/2OG. Upon oxidation, both ALKBH4/Fe(II) and the C15A/C17A/Fe(II) mutant showed the formation of an

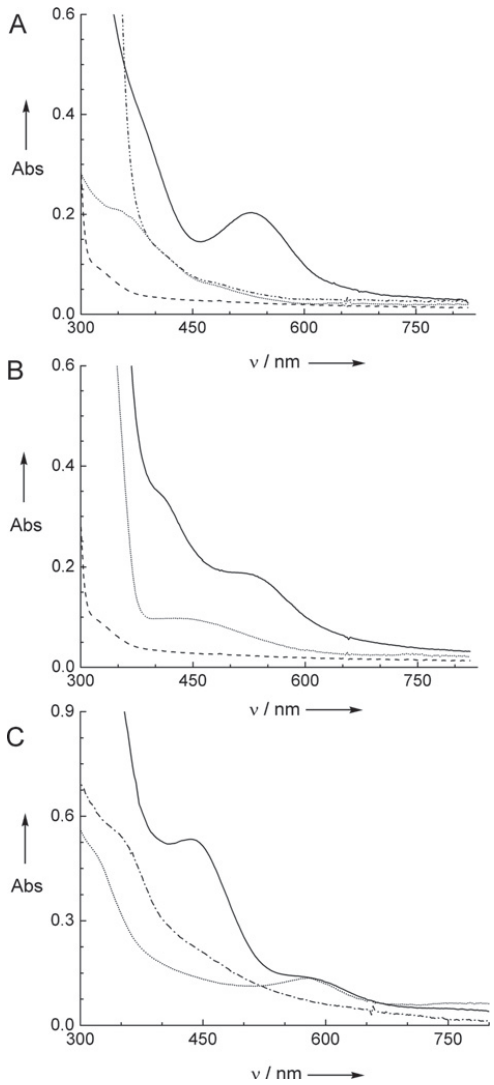


Figure 6 UV–Vis absorption spectroscopy of Fe-reconstituted ALKBH4

(**A**) Purified ALKBH4 (dashed line), ALKBH4/Fe(II) recorded under anaerobic conditions in the absence of $Na_2S_2O_4$ (dotted line), and ALKBH4/Fe(II) recorded under anaerobic conditions in the presence of $Na_2S_2O_4$ (dashed and dotted line) and after oxidation (continuous line). (**B**) ALKBH4/Fe(II)/2OG recorded under anaerobic conditions in the presence of $Na_2S_2O_4$ (dotted line) and after oxidation (continuous line). Note that the spectrum showing the optical feature of purified ALKBH4 (dashed line) from (**A**) is also included here for easier comparison. (**C**) ALKBH4/Fe(II)/2OG recorded in the presence of $Na_2S_2O_4$ after addition of NO gas under anaerobic conditions (continuous line) and after oxidation (dotted line). The corresponding spectrum of the H169A/D171A mutant [incubated with Fe(II) and 2OG] after NO exposure under anaerobic conditions is shown as a dashed line. Protein concentrations were 0.22 mM in 50 mM Hepes, pH 7.0, cooled solutions ($T \leqslant 10\,^{\circ}C$). Abs, absorbance.

absorption band centred at 530 nm ($\varepsilon \approx 920$ $M^{-1} \cdot cm^{-1}$), being accompanied by the appearance of a shoulder at higher energy ($\lambda_{max} = 370$ nm, $\varepsilon \approx 2000$ $M^{-1} \cdot cm^{-1}$) (Figure 6A, continuous line). Similarly, upon oxidation of Fe(II)- and 2OG-reconstituted ALKBH4 and C15A/C17A protein, an absorption band that was slightly blue-shifted ($\lambda_{max} = 520$ nm, $\varepsilon \approx 860$ $M^{-1} \cdot cm^{-1}$) was detected together with a shoulder at $\lambda_{max} = 404$ nm ($\varepsilon \approx 1580$ $M^{-1} \cdot cm^{-1}$) (Figure 6B, continuous line). None of the above-reported absorptions were observed in the H169A/D171A mutant, neither in the Fe-reduced nor -oxidized state. Reconstitution of ALKBH4 and the C15A/C17A protein with a 5-fold excess of Fe(II) metal ions prior to oxidation did not result in any further increase of the absorption features in the visible regions. These findings reinforce the hypothesis that (i) ALKBH4 is mononuclear in Fe content, (ii) the cysteine-rich N-terminal site does not provide an additional binding centre for the Fe(II)/Fe(III) metal ion, and furthermore (iii) an intact histidine-carboxylate site is necessary for productive Fe binding, in agreement with the results obtained in the 2OG-decarboxylation-activity assay. Since the optical features of both ALKBH4 and the C15A/C17A protein are clearly modulated by the presence of 2OG, these differences may indeed mirror modifications in the Fe co-ordination environment upon interaction with the substrate, such as the displacement of previously co-ordinated water molecules to the metal Fe. In the present study, the development of rather strong chromophores around 520–530 nm in the Fe(III) ALKBH4 and C15A/C17A protein forms closely resemble the optical spectra observed in synthetic models of non-haem Fe(III)–catecholato/phenolato complexes [40]. Therefore the spectroscopic behaviour displayed especially for the Fe(III) forms of ALKBH4 and the C15A/C17A mutant is different from that previously observed for analogous non-haem Fe proteins, such as AlkB and the Fe(II)/2OG-dependent TauD (taurine dioxygenase). AlkB exhibits an absorption band at $\lambda_{max} = 560$ nm ($\varepsilon = 1258$ $M^{-1} \cdot cm^{-1}$) [44] or, as found in other reports [12,43], a much weaker broad band at 450–500 nm, whereas TauD shows an absorption band at $\lambda_{max} = 530$ nm ($\varepsilon = 140$ $M^{-1} \cdot cm^{-1}$) [45], when both 2OG and Fe(II) are present under strict anaerobic conditions. In these proteins, exposure to oxygen in the absence of primary substrate causes the development of greenish-brown chromophores. In TauD ($\lambda_{max} = 550$ nm, $\varepsilon = 700$ $M^{-1} \cdot cm^{-1}$), the absorption envelope in the oxidized protein arises from a self-hydroxylation reaction of a tyrosine residue (Tyr^{73}) near the non-haem metal site [45]. In AlkB, the formation of a greenish-brown chromophore upon oxidation ($\lambda_{max} = 595$ nm, $\varepsilon = 960$ $M^{-1} \cdot cm^{-1}$) indicates a similar self-hydroxylation process that involves Trp^{178} [43]. In order to further study the O_2-binding properties of ALKBH4, we followed the optical changes of the Fe(II) metal core upon interaction with the dioxygen analogue NO, as performed in the EPR experiments reported earlier. Addition of NO to an anaerobic solution of ALKBH4 containing Fe(II), 2OG and $Na_2S_2O_4$ produced a strong yellowish-green chromophore, with a broad absorption band around 580 nm ($\varepsilon \approx 620$ $M^{-1} \cdot cm^{-1}$) being accompanied by a more intense peak at 436 nm ($\varepsilon \approx 2430$ $M^{-1} \cdot cm^{-1}$) (Figure 6C, continuous line). A similar spectrum was obtained for the C15A/C17A mutant (results not shown). These optical features closely resemble those reported for other non-haem Fe proteins [46] anaerobically treated with NO, such as PAH ($\lambda_{max} = 440$ nm, $\varepsilon \approx 2300$ $M^{-1} \cdot cm^{-1}$) in the presence of its co-substrate tetrahydropterin [47], which exhibits, in addition, similar EPR resonance to our NO complex. However, after injection of oxygen, the Fe(II)–NO adduct decayed into a greenish-brown chromophore ($\lambda = 580$ nm, $\varepsilon \approx 620$ $M^{-1} \cdot cm^{-1}$, and $\lambda = 770$ nm, $\varepsilon \approx 270$ $M^{-1} \cdot cm^{-1}$) (Figure 6C, dotted line), without recovery of the optical features for the Fe(III) form shown previously (Figure 6B, continuous line). This might indicate that the protein becomes more sensitive towards side

reactions upon NO binding and subsequent oxidation, since the recorded envelope in this case is very similar to the one observed in AlkB upon Trp^{178} hydroxylation [43]. Large differences can also be seen between ALKBH4 (and the C15A/C17A mutant) and the H169A/D171A mutant with respect to NO reactivity. After addition of NO to the colourless solution of the H169A/D171A/Fe(II)/2OG protein under anaerobic conditions, a yellowish solution formed without the emergence of any of the Fe–NO characteristic absorption bands, apart from a very broad absorption envelope, extending from 350 nm to 700 nm (Figure 6C, dashed and dotted line). This result supports the previous observation from the low-temperature EPR spectra (Figure 4D) that only a minor amount of Fe–NO adduct can be formed in the case of the H169A/D171A mutant, with the majority of the supplied Fe metal ions non-specifically bound to the protein.

Conclusions

In the present work, we have addressed the characteristics (electronic and magnetic) of the Fe metal core in the human AlkB homologue ALKBH4, and we have shown that an intact His^{169}-Asp^{171}-His^{254} motif is necessary for productive Fe binding and for decarboxylation activity towards 2OG. Furthermore, we have shown that the N-proximal cysteine-rich motif of ALKBH4 is not involved in Fe binding. Although the biological function of ALKBH4 remains elusive, its ability to bind and decarboxylate 2OG after metal reconstitution, together with the optical and spectroscopic fingerprints of its suggested catalytic centre, confirm for the first time that ALKBH4 truly belongs to the class of 2OG-dependent mononuclear non-haem Fe proteins. Additional work is needed to determine the nature of the transient species detected within the reoxidation process in the presence of Na_2S.

AUTHOR CONTRIBUTION

Linn Bjørnstad planned experiments, produced wild-type and mutant proteins, undertook activity measurement, analysed all of the results and wrote the manuscript. Giorgio Zoppellaro planned experiments, performed EPR and light-absorption experiments, analysed all of the results and wrote the manuscript. Ane Tomter performed EPR experiments, analysed results and wrote the manuscript. Pål Falnes designed wild-type protein and mutants, provided experience in AlkB protein, analysed all of the results and wrote the manuscript. Kristoffer Andersson planned physical chemistry experiments, provided experience in non-haem Fe proteins, analysed all of the results and wrote the manuscript.

FUNDING

This work was supported by the PEOPLE Marie Curie Actions Intra-European Fellowship within the 7th European Community Framework Programme [grant number PIEF-GA-2009-235237 (to G.Z.)], the Norwegian Cancer Society [grant number PR-2007-0132 (to L.G.B.)], the Research Council of Norway [FUGE program grant number 159013/S10 (to P.Ø.F.) and grant number 177661/V30 (to K.K.A.)] and the Polish-Norwegian Research Fund [grant number PNRF-143-AI-1/07 (to P.Ø.F.)].

REFERENCES

1 Costas, M., Mehn, M. P., Jensen, M. P. and Que, Jr, L. (2004) Dioxygen activation at mononuclear nonheme iron active sites: enzymes, models, and intermediates. Chem. Rev. **104**, 939–986
2 Schofield, C. J. and Zhang, Z. (1999) Structural and mechanistic studies on 2-oxoglutarate-dependent oxygenases and related enzymes. Curr. Opin. Struct. Biol. **9**, 722–731
3 Neidig, M. L., Brown, C. D., Light, K. M., Fujimori, D. G., Nolan, E. M., Price, J. C., Barr, E. W., Bollinger, Jr, J. M., Krebs, C., Walsh, C. T. and Solomon, E. I. (2007) CD and MCD of CytC3 and taurine dioxygenase: role of the facial triad in α-KG-dependent oxygenases. J. Am. Chem. Soc. **129**, 14224–14231
4 Price, J. C., Barr, E. W., Tirupati, B., Bollinger, Jr, J. M. and Krebs, C. (2003) The first direct characterization of a high-valent iron intermediate in the reaction of an α-ketoglutarate-dependent dioxygenase: a high-spin Fe(IV) complex in taurine/α-ketoglutarate dioxygenase (TauD) from *Escherichia coli*. Biochemistry **42**, 7497–7508
5 Que, Jr, L., and Ho, R. Y. (1996) Dioxygen activation by enzymes with mononuclear non-heme iron active sites. Chem. Rev. **96**, 2607–2624
6 Hegg, E. L. and Que, Jr, L. (1997) The 2-His-1-carboxylate facial triad – an emerging structural motif in mononuclear non-heme iron(II) enzymes. Eur. J. Biochem. **250**, 625–629
7 Neidig, M. L., Kavana, M., Moran, G. R. and Solomon, E. I. (2004) CD and MCD studies of the non-heme ferrous active site in (4-hydroxyphenyl)pyruvate dioxygenase: correlation between oxygen activation in the extradiol and α-KG-dependent dioxygenases. J. Am. Chem. Soc. **126**, 4486–4487
8 Purpero, V. and Moran, G. R. (2007) The diverse and pervasive chemistries of the α-keto acid dependent enzymes. J. Biol. Inorg. Chem. **12**, 587–601
9 Zhou, J., Gunsior, M., Bachmann, B. O., Townsend, C. A. and Solomon, E. I. (1998) Substrate binding to the α-ketoglutarate-dependent non-heme iron enzyme clavaminate synthase 2: coupling mechanism of oxidative decarboxylation and hydroxylation. J. Am. Chem. Soc. **120**, 13539–13540
10 Valegard, K., van Scheltinga, A. C., Lloyd, M. D., Hara, T., Ramaswamy, S., Perrakis, A., Thompson, A., Lee, H. J., Baldwin, J. E., Schofield, C. J. et al. (1998) Structure of a cephalosporin synthase. Nature **394**, 805–809
11 Falnes, P. O., Johansen, R. F. and Seeberg, E. (2002) AlkB-mediated oxidative demethylation reverses DNA damage in *Escherichia coli*. Nature **419**, 178–182
12 Trewick, S. C., Henshaw, T. F., Hausinger, R. P., Lindahl, T. and Sedgwick, B. (2002) Oxidative demethylation by *Escherichia coli* AlkB directly reverts DNA base damage. Nature **419**, 174–178
13 Delaney, J. C., Smeester, L., Wong, C., Frick, L. E., Taghizadeh, K., Wishnok, J. S., Drennan, C. L., Samson, L. D. and Essigmann, J. M. (2005) AlkB reverses etheno DNA lesions caused by lipid oxidation *in vitro* and *in vivo*. Nat. Struct. Mol. Biol. **12**, 855–860
14 Falnes, P. O. (2004) Repair of 3-methylthymine and 1-methylguanine lesions by bacterial and human AlkB proteins. Nucleic Acids Res. **32**, 6260–6267
15 Frick, L. E., Delaney, J. C., Wong, C., Drennan, C. L. and Essigmann, J. M. (2007) Alleviation of 1,N^6-ethanoadenine genotoxicity by the *Escherichia coli* adaptive response protein AlkB. Proc. Natl. Acad. Sci. U.S.A. **104**, 755–760
16 Koivisto, P., Robins, P., Lindahl, T. and Sedgwick, B. (2004) Demethylation of 3-methylthymine in DNA by bacterial and human DNA dioxygenases. J. Biol. Chem. **279**, 40470–40474
17 Mishina, Y., Yang, C. G. and He, C. (2005) Direct repair of the exocyclic DNA adduct 1,N^6-ethenoadenine by the DNA repair AlkB proteins. J. Am. Chem. Soc. **127**, 14594–14595
18 Falnes, P. O., Klungland, A. and Alseth, I. (2007) Repair of methyl lesions in DNA and RNA by oxidative demethylation. Neuroscience **145**, 1222–1232
19 Aravind, L. and Koonin, E. V. (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol. **2**, RESEARCH0007
20 Han, Z., Niu, T., Chang, J., Lei, X., Zhao, M., Wang, Q., Cheng, W., Wang, J., Feng, Y. and Chai, J. (2010) Crystal structure of the FTO protein reveals basis for its substrate specificity. Nature **464**, 1205–1209
21 Gerken, T., Girard, C. A., Tung, Y. C., Webby, C. J., Saudek, V., Hewitson, K. S., Yeo, G. S., McDonough, M. A., Cunliffe, S., McNeill, L. A. et al. (2007) The obesity-associated *FTO* gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science **318**, 1469–1472
22 Westbye, M. P., Feyzi, E., Aas, P. A., Vagbo, C. B., Talstad, V. A., Kavli, B., Hagen, L., Sundheim, O., Akbari, M., Liabakk, N. B. et al. (2008) Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J. Biol. Chem. **283**, 25046–25056
23 Duncan, T., Trewick, S. C., Koivisto, P., Bates, P. A., Lindahl, T. and Sedgwick, B. (2002) Reversal of DNA alkylation damage by two human dioxygenases. Proc. Natl. Acad. Sci. U.S.A. **99**, 16660–16665
24 Ringvoll, J., Moen, M. N., Nordstrand, L. M., Meira, L. B., Pang, B., Bekkelund, A., Dedon, P. C., Bjelland, S., Samson, L. D., Falnes, P. O. and Klungland, A. (2008) AlkB homologue 2-mediated repair of ethenoadenine lesions in mammalian DNA. Cancer Res. **68**, 4142–4149
25 Ringvoll, J., Nordstrand, L. M., Vagbo, C. B., Talstad, V., Reite, K., Aas, P. A., Lauritzen, K. H., Liabakk, N. B., Bjork, A., Doughty, R. W. et al. (2006) Repair deficient mice reveal mABH2 as the primary oxidative demethylase for repairing 1meA and 3meC lesions in DNA. EMBO J. **25**, 2189–2198

26 Fu, D., Brophy, J. A., Chan, C. T., Atmore, K. A., Begley, U., Paules, R. S., Dedon, P. C., Begley, T. J. and Samson, L. D. (2010) Human AlkB homolog ABH8 is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. Mol. Cell. Biol. **30**, 2449–2459

27 Fu, Y., Dai, Q., Zhang, W., Ren, J., Pan, T. and He, C. (2010) The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. Angew. Chem. Int. Ed. **49**, 1–5

28 Songe-Moller, L., van den Born, E., Leihne, V., Vagbo, C. B., Kristoffersen, T., Krokan, H. E., Kirpekar, F., Falnes, P. O. and Klungland, A. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. Mol. Cell. Biol. **30**, 1814–1827

29 Pan, Z., Sikandar, S., Witherspoon, M., Dizon, D., Nguyen, T., Benirschke, K., Wiley, C., Vrana, P. and Lipkin, S. M. (2008) Impaired placental trophoblast lineage differentiation in Alkbh1$^{-/-}$ mice. Dev. Dyn. **237**, 316–327

29a Tidwell, T. T. (2001) William Schlenk: the man behind the flask. Angew. Chem. Int. Ed. **40**, 331–337

30 Solomon, E. I. and Zhang, Y. (1992) The electronic structures of active sites in non-heme iron enzymes. Acc. Chem. Res. **25**, 343–352

31 McCleverty, J. A. (2004) Chemistry of nitric oxide relevant to biology. Chem. Rev. **104**, 403–418

32 Chen, V. J., Orville, A. M., Harpel, M. R., Frolik, C. A., Surerus, K. K., Munck, E. and Lipscomb, J. D. (1989) Spectroscopic studies of isopenicillin N synthase. A mononuclear nonheme Fe^{2+} oxidase with metal coordination sites for small molecules and substrate. J. Biol. Chem. **264**, 21677–21681

33 Arciero, D. M. and Lipscomb, J. D. (1986) Binding of ^{17}O-labeled substrate and inhibitors to protocatechuate 4,5-dioxygenase-nitrosyl complex. Evidence for direct substrate binding to the active site Fe^{2+} of extradiol dioxygenases. J. Biol. Chem. **261**, 2170–2178

34 Arciero, D. M., Orville, A. M. and Lipscomb, J. D. (1985) [^{17}O]Water and nitric oxide binding by protocatechuate 4,5-dioxygenase and catechol 2,3-dioxygenase. Evidence for binding of exogenous ligands to the active site Fe^{2+} of extradiol dioxygenases. J. Biol. Chem. **260**, 14035–14044

35 Chiou, Y.-M. and Que, Jr, L. (1995) Model studies of α-keto acid-dependent nonheme iron enzymes: nitric oxide adducts of Fe^{II}(L)(O2CCOPh)(ClO4) complexes. Inorg. Chem. **34**, 3270–3278

36 Arciero, D. M., Lipscomb, J. D., Huynh, B. H., Kent, T. A. and Munck, E. (1983) EPR and Mossbauer studies of protocatechuate 4,5-dioxygenase. Characterization of a new Fe^{2+} environment. J. Biol. Chem. **258**, 14981–14991

37 McEvoy, J. P. and Brudvig, G. W. (2008) Redox reactions of the non-heme iron in photosystem II: an EPR spectroscopic study. Biochemistry **47**, 13394–13403

38 Vass, I., Sass, L., Spetea, C., Bakou, A., Ghanotakis, D. F. and Petrouleas, V. (1996) UV-B-induced inhibition of photosystem II electron transport studied by EPR and chlorophyll fluorescence. Impairment of donor and acceptor side components. Biochemistry **35**, 8964–8973

39 Weisser, J. T., Nilges, M. J., Sever, M. J. and Wilker, J. J. (2006) EPR investigation and spectral simulations of iron-catecholate complexes and iron-peptide models of marine adhesive cross-links. Inorg. Chem. **45**, 7736–7747

40 Bruijnincx, P. C., Lutz, M., Spek, A. L., Hagen, W. R., van Koten, G. and Gebbink, R. J. (2007) Iron(III)-catecholato complexes as structural and functional models of the intradiol-cleaving catechol dioxygenases. Inorg. Chem. **46**, 8391–8402

41 Hagedoorn, P.-L., Schmidt, P. P., Andersson, K. K., Hagen, W. R., Flatmark, T. and Martínez, A. (2001) The effect of substrate, dihydrobiopterin, and dopamine on the EPR spectroscopic properties and the midpoint potential of the catalytic iron in recombinant human phenylalanine hydroxylase. J. Biol. Chem. **276**, 22850–22856

42 Martinez, A., Andersson, K. K., Haavik, J. and Flatmark, T. (1991) Recombinant human tyrosine hydroxylase isozymes. Reconstruction with iron and inhibitory effect of other metal ions. Eur. J. Biochem. **198**, 675–682

43 Henshaw, T. F., Feig, M. and Hausinger, R. P. (2004) Aberrant activity of the DNA repair enzyme AlkB. J. Inorg. Biochem. **98**, 856–861

44 Mishina, Y., Chen, L. X. and He, C. (2004) Preparation and characterization of the native iron(II)-containing DNA repair AlkB protein directly from *Escherichia coli*. J. Am. Chem. Soc. **126**, 16930–13936

45 Ryle, M. J., Liu, A., Muthukumaran, R. B., Ho, R. Y., Koehntop, K. D., McCracken, J., Que, Jr, L. and Hausinger, R. P. (2003) O_2- and α-ketoglutarate-dependent tyrosyl radical formation in TauD, an α-keto acid-dependent non-heme iron dioxygenase. Biochemistry **42**, 1854–1862

46 Orville, A. M. and Lipscomb, J. D. (1993) Simultaneous binding of nitric oxide and isotopically labeled substrates or inhibitors by reduced protocatechuate 3,4-dioxygenase. J. Biol. Chem. **268**, 8596–8607

47 Han, A. Y., Lee, A. Q. and Abu-Omar, M. M. (2006) EPR and UV–vis studies of the nitric oxide adducts of bacterial phenylalanine hydroxylase: effects of cofactor and substrate on the iron environment. Inorg. Chem. **45**, 4277–4283

48 Katoh, K., Kuma, K., Toh, H. and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic. Acids Res. **33**, 511–518

49 Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M. and Barton, G. J. (2009) Jalview Version 2 – a multiple sequence alignment editor and analysis workbench. Bioinformatics **25**, 1189–1191

50 Yu, B., Edstrom, W. C., Benach, J., Hamuro, Y., Weber, P. C., Gibney, B. R. and Hunt, J. F. (2006) Crystal structures of catalytic complexes of the oxidative DNA/RNA repair enzyme AlkB. Nature **439**, 879–884

Received 11 October 2010/14 December 2010; accepted 20 December 2010
Published as BJ Immediate Publication 20 December 2010, doi:10.1042/BJ20101667

II

OPEN ACCESS Freely available online

PLOS ONE

Human ALKBH4 Interacts with Proteins Associated with Transcription

Linn G. Bjørnstad[1], Trine J. Meza[1], Marit Otterlei[2], Solveig M. Olafsrud[3,4], Leonardo A. Meza-Zepeda[3,4], Pål Ø. Falnes[1]*

1 Department of Molecular Biosciences, University of Oslo, Oslo, Norway, **2** Department of Cancer Research and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway, **3** Genomics Core Facility, Department of Molecular Biosciences, University of Oslo, Oslo, Norway, **4** Department of Tumor Biology, the Norwegian Radium Hospital, Oslo University Hospital, Oslo, Norway

Abstract

The Fe(II)- and 2-oxoglutarate (2OG)-dependent dioxygenase AlkB from *E. coli* is a demethylase which repairs alkyl lesions in DNA, as well as RNA, through a direct reversal mechanism. Humans possess nine AlkB homologs (ALKBH1-8 and FTO). ALKBH2 and ALKBH3 display demethylase activities corresponding to that of AlkB, and both ALKBH8 and FTO are RNA modification enzymes. The biochemical functions of the rest of the homologs are still unknown. To increase our knowledge on the functions of ALKBH4 and ALKBH7 we have here performed yeast two-hybrid screens to identify interaction partners of the two proteins. While no high-confidence hits were detected in the case of ALKBH7, several proteins associated with chromatin and/or involved in transcription were found to interact with ALKBH4. For all interaction partners, the regions mediating binding to ALKBH4 comprised domains previously reported to be involved in interaction with DNA or chromatin. Furthermore, some of these partners showed nuclear co-localization with ALKBH4. However, the global gene expression pattern was only marginally altered upon ALKBH4 over-expression, and larger effects were observed in the case of ALKBH7. Although the molecular function of both proteins remains to be revealed, our findings suggest a role for ALKBH4 in regulation of gene expression or chromatin state.

Citation: Bjørnstad LG, Meza TJ, Otterlei M, Olafsrud SM, Meza-Zepeda LA, et al. (2012) Human ALKBH4 Interacts with Proteins Associated with Transcription. PLoS ONE 7(11): e49045. doi:10.1371/journal.pone.0049045

Editor: Ben C.B. Ko, Chinese University of Hong Kong, Hong Kong

Received July 13, 2012; **Accepted** October 3, 2012; **Published** November 8, 2012

Funding: This work was supported by the Norwegian Cancer Society (www.kreftforeningen.no, grant number PR-2007-0132), the Research Council of Norway (www.forskningsradet.no, FUGE program grant number 159013/S10) and the Polish-Norwegian Research Fund (www.fbn.opi.org.pl, grant number PNRF-143-AI-1/07). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pal.falnes@imbv.uio.no

Introduction

The superfamily of Fe(II)- and 2-oxoglutarate (2OG)-dependent dioxygenases comprise enzymes which catalyze oxidation reactions in a diverse set of biological processes such as post-translational modification of collagen, the hypoxic response pathway and epigenetic regulation [1–6]. These proteins are characterized by their catalytic requirement for ferrous iron as well as the co-substrate 2OG. The primary oxidation reactions they catalyze are coupled to decarboxylation of 2OG, yielding succinate and CO_2. The *E. coli* AlkB protein [7] is an Fe(II)/2OG dioxygenase involved in DNA and RNA repair, and is induced as part of the adaptive response to alkylation damage. Targeting alkyl lesions at N1-position in purines and N3-position in pyrimidines, AlkB directly reverses the base damage by an oxidative mechanism that involves hydroxylation of the alkyl group, which is consequently destabilized and spontaneously released [8,9]. The repertoire of AlkB substrates has extended from the originally identified simple methyl lesions 1-methyladenine (1-meA) and 3-methylcytosine (1-meC), to also comprise larger adducts, such as ethyl, propyl and etheno groups [10–13], as well as methylated RNA [14].

In mammals, nine AlkB homologs have been reported; ALKBH1-8, as well as the fat mass and obesity protein (FTO) [15,16]. ALKBH2 and ALKBH3 have similar activities as AlkB, ALKBH2 being most active on dsDNA, while ALKBH3 preferentially demethylates ssDNA and ssRNA [14]. FTO has also, with its weak activity towards 3-methylthymine (3-meT) in ssDNA and 3-methyluracil (3-meU) in ssRNA, been implicated in nucleic acid repair [16,17], but the recent identification of N^6-methyladenosine (6-meA) in ssRNA as a preferred substrate indicates a role for FTO in regulating mRNA modification [18]. With the recent demonstration of ALKBH8 being involved in hypermodification of tRNA wobble uridines [19–22], the function of the mammalian ALKBH proteins was definitely shown to extend beyond nucleic acid repair. ALKBH1, the mammalian homolog with highest similarity to AlkB, has also been reported to display DNA repair activity [23]. However, the significance of this activity, which has not been confirmed by others, is unclear, and there are indications rather pointing towards a function for ALKBH1 in epigenetic gene regulation, potentially through histone demethylation [24,25], thereby supporting the suggestion that some of the human ALKBHs are involved in protein demethylation [26,27]. In line with this, the *S. pombe* AlkB homolog Ofd2 was recently reported to interact with histones [28]. While 2-oxoglutarate decarboxylase activity has been demonstrated for ALKBH4 and ALKBH5 [29,30], their primary substrates and biological functions still remain, together with those of ALKBH1, ALKBH6 and ALKBH7, to be revealed.

The present study focuses on the human ALKBH4 and ALKBH7 proteins. Through yeast two-hybrid screening, we show binding of ALKBH4 to several proteins with associations to chromatin regulation and transcription. Furthermore, for a selection of these, we demonstrate nuclear co-localization with ALKBH4. However, as revealed through gene expression profiling, ALKBH4 over-expression marginally affects the expression pattern in HEK293 cells, while over-expression of ALKBH7 influences biological pathways such as cell cycle, DNA repair and spermatogenesis, and positively regulates a number of genes involved in meiotic recombination. Thus, this work provides novel insight into the biological function of mammalian AlkB homologs for which no biochemical activity has yet been reported.

Results

Yeast Two-hybrid Screens Identified ALKBH4 Binding Partners Involved in Transcription

To improve our knowledge on ALKBH4 and ALKBH7 function, we searched to identify interaction partners of these proteins through yeast two-hybrid (Y2H) screens. Screens were performed using ALKBH4 as bait against two different human libraries, one from placenta and another from fetal brain (Hybrigenics, France). ALKBH7 was used as bait to screen the fetal brain library. While we did not obtain any hits considered highly confident in the ALKBH7 screen, a total of ten such hits were detected in the two screens concerning ALKBH4 (Table 1). Notably, five of these proteins have been associated with transcription and chromatin modification, suggesting a function of ALKBH4 in gene regulation. Of these, the transcriptional co-activator and histone acetyltransferase (HAT) p300 [31] was the only protein identified in both screens and, of note, the highly similar p300 paralog CBP was not detected in any of them. The homeotic transcription factor ATBF1 [32,33] and the tissue specific heat-shock transcription factor HSF4 [34,35] were identified as ALKBH4 partners exclusively in the placenta screen. The proteins AF9 and ENL, which have similar biological functions and display very high sequence homology (56% identity throughout the entire sequence) [36] were identified in the placenta and brain screens, respectively. The specific functions of AF9 and ENL are currently unknown, but they have both been associated with histone modification and transcriptional elongation [37–39].

To further investigate whether the observed interactions are dependent on the enzymatic activity of ALKBH4, an enzymatically inactive mutant ($ALKBH4^{H169A/D171A}$) was used as bait in a Y2H screen against the placenta library. Notably, this screen retrieved a very similar set of proteins as the screen performed with wild-type ALKBH4 (Table 1), indicating that the observed interactions can occur independently of the oxygenase activity of ALKBH4.

ALKBH4 Binding is Mediated by Chromatin-associated Domains

The clones identified in the Y2H screens usually represented non-full-length fragments of the interactants. The overlapping sequences present in all clones representing the same partner protein define the part of the protein that is responsible for the interaction, the so-called selected interaction domain (SID). A schematic representation of the transcription/chromatin-related high confidence hits, depicting the SID involved in the interaction with ALKBH4, is shown in Figure 1. Interestingly, for all five proteins the SID encompassed only a limited region, and in all cases this region included annotated domains reported to interact with DNA and/or chromatin. For both AF9 and ENL, the ALKBH4-interacting part mapped to the chromatin-associated YEATS domain, while the portion of ATBF1 that binds ALKBH4 was shown to encompass two of the numerous C_2H_2-type zinc fingers found in this protein. Moreover, we found the interaction between ALKBH4 and HSF4 to be mediated through the amino-terminal DNA binding domain (DBD) of the latter. Finally, the region of p300 mediating the ALKBH4-interaction covered both the bromodomain and plant homeodomain (PHD).

Since this region is adjacent to the HAT domain and since bromo domain are known to bind to acetylated proteins, we considered the possibility that ALKBH4 is acetylated by p300. To address this issue, we used an anti-FLAG antibody to immunoprecipitate FLAG-ALKBH4 from cell extracts containing over-expressed p300 as well as FLAG-ALKBH4. However, no *in vivo* ALKBH4 acetylation was detected (data not shown).

ALKBH4 Co-localizes with the Transcriptional Proteins AF9, ENL and p300 in the Nucleoplasm and Nucleoli

We have made extensive efforts to verify the interactions that were observed in the Y2H screen by independent methods, such as pull-down experiments with recombinant GST-tagged proteins and co-immunoprecipitation of tagged, co-expressed proteins from cell lysates (experimental outlines can be found in the Materials and Methods section). However, none of these efforts were successful with respect to robust verification, possibly because the interactions are transient and of low affinity. Nonetheless, given the striking overrepresentation of DNA/chromatin binding moieties of transcription-associated proteins among the interactions, we still considered it likely that the corresponding interactions are biologically relevant and occurring in mammalian cells *in vivo*. To investigate this, the chromatin-related proteins that were detected in both screens, AF9, ENL and p300, were selected for co-localization studies with ALKBH4. HeLa cells were transiently co-transfected with a plasmid encoding ALKBH4 fused to either Enhanced Cyan Fluorescent Protein (ALKBH4-ECFP) or Enhanced Yellow Fluorescent Protein (ALKBH4-EYFP) in combination with plasmids encoding EYFP/ECFP-fusions of the selected partner proteins or truncations of these, and localization patterns were analyzed by confocal microscopy. First, we determined the subcellular localization of ALKBH4 (ALKBH4-EYFP) alone, which, consistent with a previous report [40], was found to be localized both in the nucleus and the cytoplasm, with similarly strong signals detected in the two compartments (Figure 2A).

Co-localization with the AF9/ENL YEATS domain. ALKBH4 showed nuclear co-localization with both AF9 and ENL in which partial overlap in distinct spots in nucleoplasm and nucleoli was observed for both proteins (Figure 2B and 2C and data not shown). In order to address the significance of the YEATS domain for these presumable interactions, we examined the co-localization patterns of ALKBH4 with two truncated ENL versions, which expressed either the YEATS domain (EYFP-ENL_{YEATS}) or the C-terminal part (ENL_C-EYFP). A co-localization pattern similar to that of full-length ENL was obtained for the YEATS only version (Figure 2D). However, deletion of the YEATS domain strongly reduced the co-localization of ENL with ALKBH4 in nucleoli (Figure 2E). This supports the results from the Y2H screens, strengthening the indication of an interaction between ALKBH4 and AF9/ENL, with the YEATS domain of latter being important for mediating the interaction.

Co-localization with the bromodomain and PHD finger of p300. In our Y2H screens, the p300 clones identified to bind

Table 1. High confidence[a] ALKBH4 interacting proteins identified by the yeast two-hybrid system.

Cellular process	Gene	Length (aa)	SID[b]	Domain(s) encompassed by SID	Screen		
					Fetal brain	Placenta	Placenta ($ALKBH4^{H169A/D171A}$)
Transcription	*ATBF1*	3703	1545–1626	C_2H_2 zinc fingers	–	+	+
	AF9	568	10–151	YEATS	+	–	–
	ENL	559	4–151	YEATS	–	+	(+)[c]
	HSF4	462	1–93	DBD	–	+	+
	p300	2414	1054–1352	Bromodomain, PHD	+	+	+
Other	*TES*	421	1–375	PET, LIM1–2	–	+	+
	EIF3C	914	749–882	PCI	+[d]	+	–
	MTMR6	621	251–539	PTP	+	(+)[c]	–
	PSMA6	246	11–232	Proteasome	+	–	–
	GID: 13396337	242	13–66	Transmembrane domain	+	–	–

[a]PBS (predicted biological score) of A or B. PBS (calculated according to [75]) represents the probability of an interaction to be non-specific, and refers to an e-value with defined thresholds to rank the results in the high-to-low confidence categories A–D.
[b]SID (selected interaction domain) refers to the amino acid region shared by all prey fragments matching the same reference protein,
[c](+) Low confidence hit (PBS D),
[d]EIF3CL. Grey shade, transcription related interactants. YEATS, Yaf9 ENL AF9 TAF14 Sas5; DBD, DNA binding domain; PHD, plant homeodomain; PET, Prickle Espinas Testin; LIM, Lin-11 Isl-1 Mec-3; PCI, Proteasome, COP9, Initiation factor-3; PTP, protein tyrosine phosphatase.
doi:10.1371/journal.pone.0049045.t001

ALKBH4 all encompassed both the bromodomain and PHD finger, a protein domain combination that has been shown to compose a functional protein moiety [41]. Thus, in order to determine if p300 co-localizes with ALKBH4, we decided to use an EYFP-tagged fragment of p300 covering the region of the bromodomain and PHD only (aa 1039–1285, EYFP-$p300_{BP}$), instead of the full-length protein. Indeed, similar to what was observed for AF9 and ENL, $p300_{BP}$ co-localized with ALKBH4 in spots in nucleoli (Figure 2F). A similar nuclear localization pattern and nucleolar presence has been previously reported for over-expressed, full-length p300 [42]. Our results thus indicate an interaction between ALKBH4 and $p300_{BP}$. However, the in-

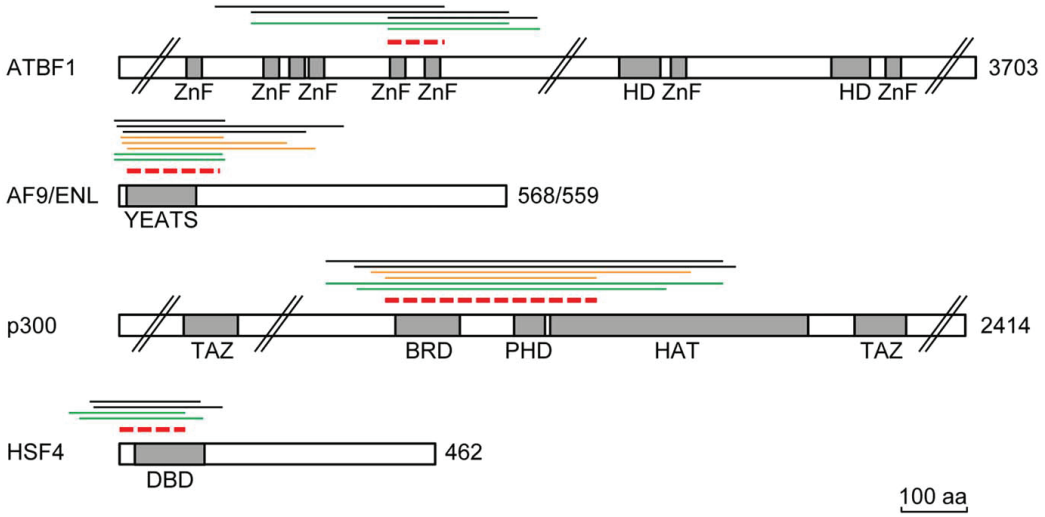


Figure 1. Schematic representation of yeast two-hybrid high confidence hits involved in transcription. Individual prey fragment clones and the resulting selected interaction domains (SIDs) reported to bind ALKBH4 are indicated above each protein; black lines, placenta library; orange lines, fetal brain library; green lines, placenta library screened with $ALKBH4^{H169A/D171A}$; red dashed lines, SIDs. Grey boxes indicate protein domains. Proteins and domains are drawn to scale according to the InterPro (version 4.8) and PROSITE (release 20.68) databases [73,74].//indicates regions omitted for simplicity. ZnF, C_2H_2 zinc finger; HD, homeodomain; YEATS, Yaf9 ENL AF9 Taf14 Sas5; TAZ, transcription adaptor putative zinc finger; BRD, bromodomain; PHD plant homeodomain; HAT histone acetyl transferase; DBD, DNA binding domain. Bar, 100 aa.
doi:10.1371/journal.pone.0049045.g001

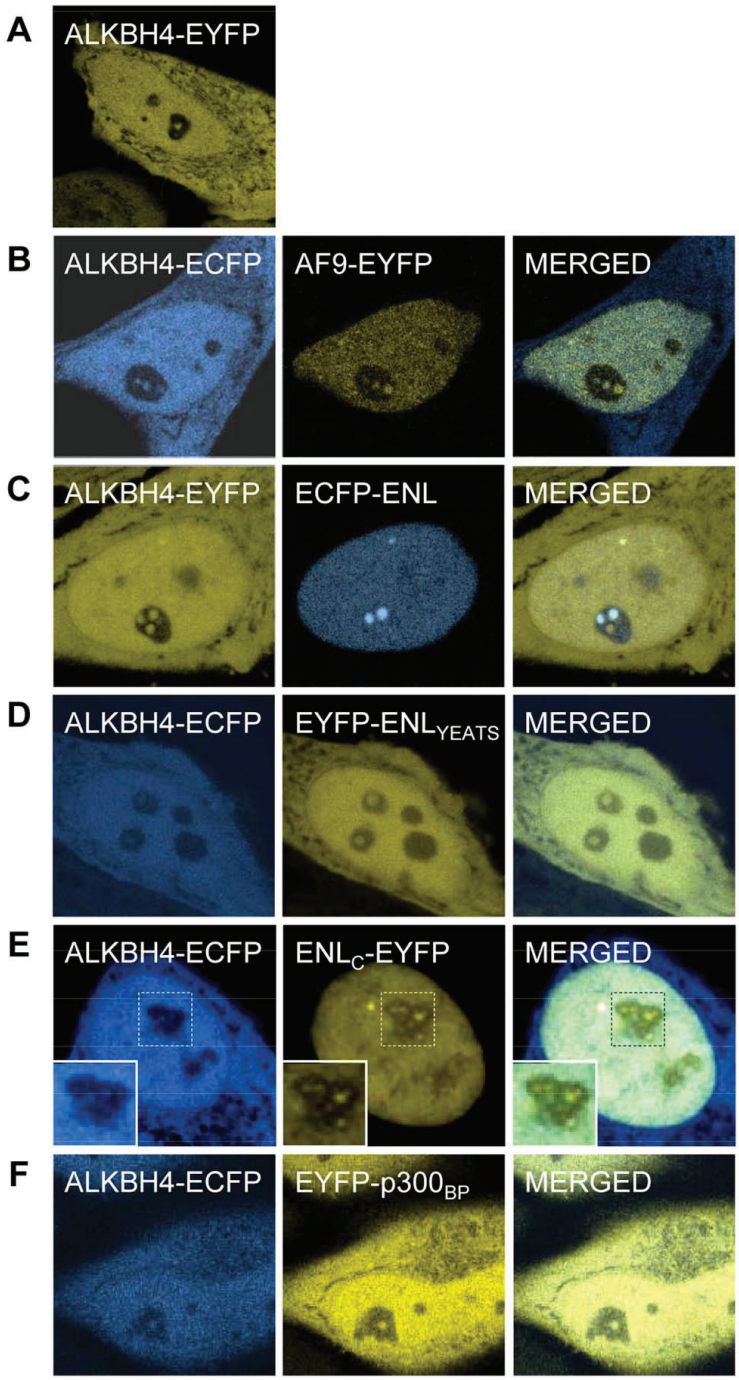

A
ALKBH4-EYFP
B
ALKBH4-ECFP
AF9-EYFP
MERGED
C
ALKBH4-EYFP
ECFP-ENL
MERGED
D
ALKBH4-ECFP
EYFP-ENL$_{YEATS}$
MERGED
E
ALKBH4-ECFP
ENL$_{C}$-EYFP
MERGED
F
ALKBH4-ECFP
EYFP-p300$_{BP}$
MERGED

Figure 2. Subcellular localization of ALKBH4 and its co-localization with transcription-associated proteins. (**A**) Subcellular localization of ALKBH4-EYFP in HeLa cells. Co-expression of (**B**) ALKBH4-ECFP and AF9-EYFP, (**C**) ALKBH4-EYFP and ECFP-ENL, (**D**) ALKBH4-ECFP and EYFP-ENL_{YEATS}, (**E**) ALKBH4-ECFP and ENL_{C}-EYFP and (**F**) ALKBH4-ECFP and EYFP-$p300_{BP}$, as analyzed by confocal fluorescence microscopy. Insets are enlargements of boxed areas.
doi:10.1371/journal.pone.0049045.g002

dividual contribution of the two domains considered here can not be determined from this experiment.

ALKBH4 Partly Co-localizes with the RNA Polymerase I Complex

Being the site of ribosome biogenesis, the nucleolus also comprises the process of ribosomal RNA (rRNA) synthesis. Our observations of ALKBH4 co-localizing in nucleolar foci with all three transcription-associated proteins examined made us speculate if ALKBH4 could potentially have a function in transcription of ribosomal DNA (rDNA) genes. We therefore examined whether ALKBH4 and ENL co-localize with the RNA polymerase I subunit RPA43. Thus, HeLa cells were transiently co-transfected with plasmids encoding ALKBH4-EYFP, ECFP-ENL and a Red Fluorescent Protein (RFP)-fusion of RPA43 (RPA43-RFP). RPA43 was observed at the surroundings of the distinct nucleolar ALKBH4/ENL foci with partially overlapping staining observed (Figure S1, insert and nucleolus at the lower right). Partial co-localization of ALKBH4 with ENL in nucleoplasmic and nucleolar spots was detected, as mentioned above. Notably, we observed a larger amount of ENL foci overlapping with RPA43 foci compared to the overlap between ALKBH4 and RPA43. Hence, these results suggest that ALKBH4 is not primarily involved in nucleolar RNA polymerase I-dependent rDNA transcription.

Effects of ALKBH4 and ALKBH7 on Global Gene Expression and DNA Methylation Patterns

We further considered the possibility that ALKBH4 may itself be capable of modulating transcription. To investigate this, we generated a stable HEK293 cell line in which the ALKBH4 encoding gene was introduced at a specific, transcriptionally active, genomic locus, behind a tetracycline-inducible promoter, thereby enabling controlled over-expression of ALKBH4. Ectopic ALKBH4 expression was induced upon treatment with the tetracycline analog doxycycline (DOX), and the resulting ALKBH4 increase was verified on the mRNA and protein levels, by qPCR and Western blotting, respectively (Figures 3A and B). The global gene expression profiles of the ALKBH4 over-expressing cell line and the non-induced, parental cell line were subsequently compared in a microarray-based genome-wide expression analysis. Surprisingly, very small effects on gene expression were observed upon over-expression of ALKBH4. Actually, none of the genes were up- or down-regulated above 2-fold, and only 22 genes showed differential expression when the fold change stringency was reduced to 1.35 (Figure 3C). A list of differentially expressed genes (q-value <5, fold change (FC) >1.35) upon ALKBH4 over-expression is shown in Table 2. Noteworthy, the absence of *ALKBH4* itself among the up-regulated genes in the over-expressing cell line is explained by annealing of the probe to the untranslated region (UTR) of the gene (data not shown). The low number of genes affected by ALKBH4 over-expression indicates that ALKBH4 does not affect transcription at the global level in the HEK293 cells.

Furthermore, to characterize the biological functions of the genes affected by ectopic ALKBH4 expression, a gene ontology (GO) analysis was performed using MetaCore GeneGo Pathways Analysis software (GeneGo Inc.). However, the differentially expressed genes were not enriched in any molecular pathway.

Several mammalian AlkB homologs have been shown to be involved in DNA/RNA transactions, and these are basic proteins with a high pI value, thus allowing binding to the negatively charged nucleic acid backbone. In contrast, ALKBH1, ALKBH4 and ALKBH7 are acidic proteins, and it has been proposed that they may be involved in demethylating histones or other proteins (Sedgwick *et al.* 2007). Indeed, genetic ablation of ALKBH1 has been shown to dysregulate a number of genes in the mouse placenta [24], and we were therefore also interested in investigating the effects of ALKBH7 on gene regulation. Using the same approach as for ALKBH4, we generated a stable HEK293 cell line with tetracycline-inducible over-expression of ALKBH7, which was used to determine the effects of ectopic ALKBH7 levels on gene expression. Considerably stronger effects were detected in the case of ALKBH7, compared to ALKBH4. A total of 532 genes, excluding *ALKBH7*, were differentially expressed (q-value <5, FC >2) between the parental and ALKBH7 over-expressing cell lines. Of these, 197 genes were up-regulated and 335 genes were down-regulated (Figure 3C). Notably, ALKBH7 over-expression resulted in a higher number of down-regulated compared to up-regulated genes and, moreover, the most down-regulated genes were more strongly affected ($3.8<$ FC <5.4) than the most up-regulated ones ($2.9<$ FC <3.4). The most up- or down-regulated genes are listed in Table 3.

In order to determine possible enrichment of ALKBH7 affected genes in certain biological pathways, we further performed a GO analysis similar to that of ALKBH4. Twelve GO categories were determined to be significantly overrepresented (false discovery rate (FDR) <0.05) among the differentially expressed genes (Figure 3D). The pathway mostly enriched with ALKBH7 affected genes was "cell cycle and its regulation", in which 33 of 444 genes displayed expression level alterations upon ectopic ALKBH7 expression. While the great majority of these (31 genes) showed decreased expression, only two genes, *RBX1 (RING-box protein 1)* and *PRKAR1A (PKA-regulatory subunit 1A)*, were up-regulated.

ALKBH7 is annotated in the NCBI database as spermatogenesis-associated protein 11 (*SPATA11*) or spermatogenesis cell proliferation-related protein. Interestingly, the results of our GeneGo pathway analysis included "spermatogenesis" as a category of enriched genes (Figure 3D). Herein, three genes (*CREB1, PRKAR1A and GNAS1*) of a total of 22, belonging to the subcategory "transcription – CREM signaling in testis", were differentially expressed after ALKBH7 over-expression. Moreover, among the top twenty up-regulated genes, three genes involved in the process of meiotic recombination during gametogenesis were identified: *disrupted meiotic cDNA 1 (DMC1)* [43], *decreased sperm survival 1 (DSS1)* [44] and *male-specific lethal 3-like 1 (MSL3L1)* [45]. Thus, our results are consistent with an association of the *ALKBH7* gene with spermatogenesis.

To complement the gene expression data, we also performed microarray-based CpG methylation profiling in cells over-expressing either ALKBH4 or ALKBH7, as well as in the respective non over-expressing cells. In case of both proteins, the methylation pattern remained strikingly constant upon over-expression ($r^2 = 0.9975$ and 0.9916 for ALKBH4 and ALKBH7, respectively. Figure S2), thus excluding a role for ALKBH4 and ALKBH7 in regulation of global CpG methylation.

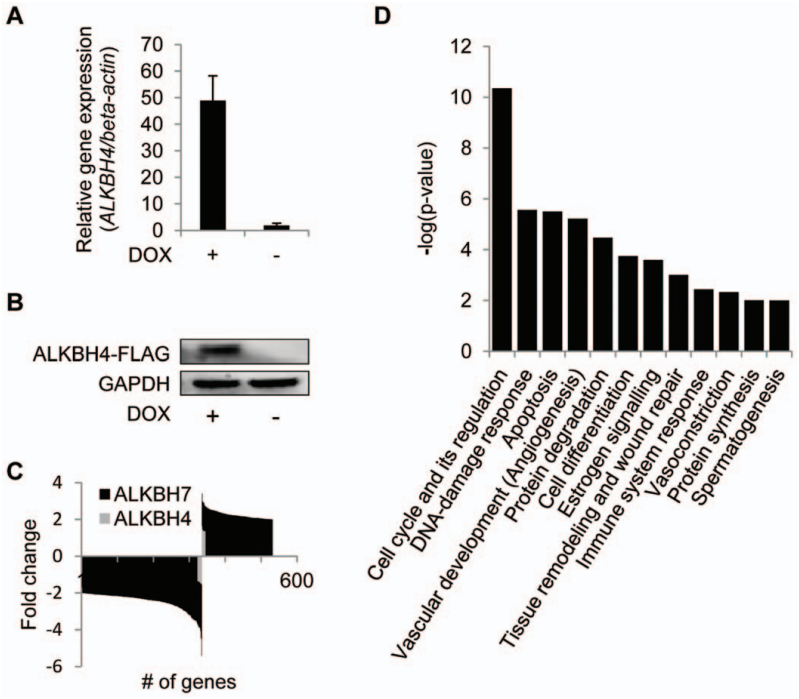


Figure 3. Effects of ectopic expression of ALKBH4 or ALKBH7 on global gene expression. (**A**) Quantitative RT-PCR analysis of relative *ALKBH4* levels in HEK293 cells stably transfected with a construct for DOX-inducible over-expression of ALKBH4-FLAG, either treated with DOX (2 μg/ml) or untreated. Results are presented as mean fold change of three independent replicates normalized to *β-actin* ± S.D. (**B**) Ectopic ALKBH4 protein levels in DOX-induced and non-induced cells, as determined by Western blot analysis. Ectopic ALKBH4 was detected using an antibody against the FLAG-tag introduced at the C-terminus of ALKBH4. GAPDH expression levels are included as loading control. (**C**) Microarray analysis of gene expression in cells over-expressing either ALKBH4 or ALKBH7 vs. non-overexpressing cells. The number of genes whose expression is altered at least 2.0-fold (ALKBH7) or 1.35-fold (ALKBH4) is indicated (**D**) MetaCore (GeneGo Inc.) analysis of molecular pathways significantly (False discovery rate (FDR) <0.05) enriched with genes affected by ectopic ALKBH7 expression. DOX, doxycycline.
doi:10.1371/journal.pone.0049045.g003

In vivo H3K79 Methylation Levels are not Affected by Ectopic ALKBH4 Expression

Regulation of transcription depends on dynamic chromatin modifications, such as histone methylation, involving proteins with either methyltransferase or demethylase activity. In contrast to the situation for methylated lysine residues found at the flexible histone tails, no demethylase has been found to reverse the activating methyl mark at the lysine 79 (K79) residue in the globular domain of H3, which in mammals is introduced by the enzyme DOT1L. However, there are indications of the existence of such enzyme activity [46,47], and a study in which treatment of cells with 2-hydroxyglutarate (2HG), an inhibitor of Fe(II)/2OG-dependent oxygenases, resulted in increased H3K79 dimethyl levels has further suggested a member of this protein family to be the enzyme responsible for H3K79 demethylation [48]. Moreover, ALKBH4 has previously been suggested to possess protein demethylase activity [26], and our results imply a role for ALKBH4 in chromatin regulation. Since we found ALKBH4 to interact with the DOT1L-associated proteins AF9 and ENL, we speculated that ALKBH4 could potentially function as a demethylase with specificity for methylated H3K79. In order to address this issue, we isolated histones from the DOX-inducible ALKBH4 over-expressing HEK293 cell line before and after DOX induction, as well as from a similarly generated cell line possessing the ability of stable, inducible over-expression of an enzymatically inactive ALKBH4 mutant ($ALKBH4^{H169A/D171A}$). Subsequently, the methylation status at the H3K79 position in the histones was determined by Western blotting. However, we did not observe any effect of ectopic ALKBH4 expression on the methylation status of histone H3K79, as similar levels of mono-, di- and trimethylated H3K79 were detected in the non-overexpressing cells as well as in those over-expressing $ALKBH4^{H169A/D171A}$ (Figure 4). This suggests that ALKBH4, if involved in demethylation of this histone residue, might be restricted to certain, presently unknown, conditions to be functional.

Discussion

In the present work, we report that proteins involved in transcription were strongly over-represented among interactants of the human oxygenase ALKBH4 identified by Y2H screens, while no convincing partners were detected for the related ALKBH7 protein. Interestingly, the regions of these transcription-associated interactants that was responsible for interaction with ALKBH4, in

Table 2. Differentially expressed genes (q-value <5, fold change >1.35) identified in ALKBH4 over-expressing cells compared to parental non over-expressing cells.

	Gene	Description	Fold Change
Up-regulated[1]	*HSPA1B*	Heat shock 70kDa protein 1B	1.63
	INSIG1	Insulin induced gene 1, transcript variant 2	1.43
	LAMA5	Laminin, alpha 5	1.43
	FAM38A	Family with sequence similarity 38, member A	1.39
	HSPA8	Heat shock 70kDa protein 8, transcript variant 1	1.38
	LOC642031	Hypothetical protein LOC642031	1.38
	FASN	Fatty acid synthase	1.36
	LOC23117	KIAA0220-like protein, transcript variant 16	1.35
	INTS1	Integrator complex subunit 1	1.35
	SEC16A	SEC16 homolog A	1.35
Down-regulated	*DDIT4*	DNA-damage-inducible transcript 4	−1.55
	LGALS1	Lectin, galactoside-binding, soluble, 1	−1.50
	LETMD1	LETM1 domain containing 1, transcript variant 1	−1.48
	LOC653994	Similar to Eukaryotic translation initiation factor 4H, transcript variant 2	−1.48
	SLC3A2	Solute carrier family 3, member 2, transcript variant 6	−1.45
	CTH	Cystathionase (cystathionine gamma-lyase), transcript variant 1	−1.43
	TSC22D3	TSC22 domain family, member 3, transcript variant 2	−1.43
	MCM7	Minichromosome maintenance complex component 7, transcript variant 1	−1.38
	DDIT3	DNA-damage-inducible transcript 3	−1.37
	STC2	Stanniocalcin 2	−1.37
	SNX5	Sorting nexin 5, transcript variant 1	−1.37
	RASSF1	Ras association (RalGDS/AF-6) domain family member 1, transcript variant C	−1.36

[1]The *ALKBH4* probe was not detected due to annealing to the UTR of the gene. Sorted by fold change.

doi:10.1371/journal.pone.0049045.t002

all cases encompassed domains which are involved in interaction with DNA and chromatin, and for some of these interactants we observed a co-localization with ALKBH4 in distinct foci in the nucleus. We also performed a global analysis of gene expression and CpG methylation changes induced by ectopic over-expression of ALKBH4 and ALKBH7. While we saw rather small effects of ALKBH4 on both expression and methylation, larger effects on gene expression were observed in the case of ALKBH7.

Protein Partners Suggest a Role for ALKBH4 in Gene Regulation

Our identification of the transcriptional co-activator p300 as an ALKBH4 partner is indicative of a gene regulatory role of ALKBH4. The region of p300 mediating the ALKBH4-interaction covered both the bromodomain and plant homeodomain (PHD) of this protein. Both domains represent motifs found in chromatin modification effector proteins, thus functioning in recruitment of remodeling-complexes to chromatin. While bromodomains recognize acetylated histones [49], the zinc-coordinating PHD finger is generally involved in binding methylated or unmodified histone H3 [50]. However, PHD fingers are often found close to bromo- or chromodomains [51], and several reports have shown a combinatorial function of the two domains [52,53]. The function of the p300 PHD finger is still to be elucidated, although it has been reported to be required, together with the bromodomain, for *in vitro* recognition of acetylated nucleosomes [41]. Notably, the histone acetyl transferase CBP, which is highly similar to p300, was not identified as an ALKBH4 interaction partner in the yeast two-hybrid screen. However, p300 HAT activity does not, in contrast to that of CBP, depend on its PHD finger, implying non-redundant functions for these co-activators [54].

Further supporting an involvement of ALKBH4 in gene regulation, we also detected the homologs AF9 and ENL as interacting partners. In addition to a hydrophobic, transcription-associated C-terminus both proteins contain an N-terminal YEATS domain [37], which we here found to mediate the ALKBH4 interaction. The YEATS domain has been named after proteins that carry such a domain (Yaf9, ENL, AF9, Taf14 and Sas5), and many of these proteins are components in transcriptional or chromatin-modifying complexes. The specific function of YEATS has not yet been determined, but several findings indicate a role in chromatin binding. The interaction between ENL and histones H1 and H3 has previously been demonstrated to be mediated through YEATS [55], and recently, AF9/ENL YEATS was found to be the module responsible for recruitment of the super elongation complex (SEC) to chromatin and the elongating Polymerase II [56]. Moreover, as a result of the first three-dimensional YEATS structure, this domain was recently suggested to provide an additional reader module of chromatin, analogous to the bromodomain and PHD finger [57,58], with the Yaf9 YEATS domain suggested to potentially bind acetylated lysine-residues in histones [57].

The interactions of ALKBH4 with the transcription factors HSF4 and ATBF1 were found to involve domains with the ability of DNA binding, these being the DNA binding domain (DBD) of

HSF4 and two of the multiple C_2H_2-type zinc fingers of ATBF1. The versatile C_2H_2 motif, which is frequently found in gene regulators, has been demonstrated to mediate interactions with proteins as well as both DNA and RNA (reviewed in [59]). Previously, two of the ATBF1 zinc fingers have been reported to bind protein [60].

Interestingly, three of the transcription-associated ALKBH4 partners, AF9, ENL and p300, have all been reported to be fused to the histone methyltransferase MLL in mixed-lineage leukemia (MLL) [37,61,62], which is characterized by abberant H3K79 dimethylation profiles [63]. While the ALKBH4 interacting YEATS domain of AF9/ENL is not present in the fusion proteins MLL-AF9 and MLL-ENL [37], the less frequent MLL-p300 fusions, which are thought to promote leukemogenesis through aberrant histone acetylation rather than methylation [61,64], retain the ALKBH4 interacting p300 region.

Molecular Function of ALKBH4

Although the present work suggests that ALKBH4 may be involved in processes such as chromatin regulation and transcription, its molecular function remains an enigma. Interestingly, two of the herein identified ALKBH4 interacting proteins with associations to chromatin regulation and transcription, AF9 and ENL, also interact with DOT1L, the enzyme responsible for methylation of lysine 79 at histone H3 (H3K79). While no demethylase with specificity for this residue has been reported so far, there are indications of such activity [46,47]. The responsible enzyme has been suggested to be an Fe(II)/2OG dioxygenase, as elevated H3K79 dimethyl levels has been observed as a result of Fe(II)/2OG dioxygenase inhibition [48]. Intriguingly, plants lack both H3K79 methylation and DOT1L orthologs [65], as well as ALKBH4. Thus, ALKBH4 could be imagined to function as an H3K79 demethylase. Consequently, we investigated whether ectopic ALKBH4 expression affected the cellular H3K79 methylation status in HEK293 cells. However, we did not detect alterations in the levels of neither mono-, di- nor trimethylated H3K79 after ALKBH4 over-expression. Related Fe(II)/2OG dioxygenase activities include JumonjiC (JmjC) protein mediated histone demethylation [2], structural and gene regulatory protein hydroxylation by prolyl-4-hydroxylase (P4H) and prolyl hydroxylase domain containing (PHD) 1–3, respectively [5,66], as well as epigenetic DNA hydroxylation by the TET enzyme family, which was recently reported to convert 5-methylcytosine (5-meC) to 5-hydroxymethylcytosine (5-hmC) [1]. Previously, the low pI-value of ALKBH4 has suggested this enzyme to possess a protein, rather than a nucleic acid substrate [26], and, consistently, no activity towards the classical AlkB substrates 1-meA and 3-meC has been detected [67]. Furthermore, we previously showed uncoupled decarboxylase activity for ALKBH4, indicating that this protein is a *bona fide* Fe(II)/2OG dioxygenase [30].Thus, a demethylase or hydroxylase activity targeting proteins or, less likely, nucleic acids, could be envisioned for ALKBH4.

Association of Genes Affected by ALKBH7 Over-expression with DNA Recombination

The previous annotation of *ALKBH7* being associated with spermatogenesis is consistent with the data presented here, as functional annotation analysis revealed spermatogenesis as one of the biological pathways enriched with genes differentially expressed upon ectopic ALKBH7 expression. Meiotic recombination is one of the central events occurring during spermatogenesis, and, consistently, three genes with functions in meiotic recombination were found among the twenty most up-regulated ones. This, in combination with the over-representation of genes related to cell cycle regulation and the DNA damage response among the differentially expressed genes, may indicate a possible role for ALKBH7 in the process of homologous recombination repair (HRR) of double strand breaks (DSBs) introduced by DNA

Table 3. The genes most differentially expressed in either direction identified in ALKBH7 over-expressing cells compared to parental non over-expressing cells.

	Gene	Description	Fold Change
Up-regulated	*ALKBH7*	AlkB, alkylation repair homolog 7	13.54
	MSL3L1	Male-specific lethal 3-like 1, transcript variant 1	3.41
	LOC643287	Similar to prothymosin alpha, transcript variant 1 (LOC643287)	2.97
	RPL21	Ribosomal protein L21	2.96
	NBPF20	Neuroblastoma breakpoint family, member 20	2.95
	PEBP1	Phosphatidylethanolamine binding protein 1	2.90
	GNL3L	Guanine nucleotide binding protein-like 3 (nucleolar)-like	2.89
	SHFM1	Split hand/foot malformation (ectrodactyly) type 1	2.85
Down-regulated	*LOC650215*	Similar to Exportin-T (tRNA exportin)	-5.42
	PKMYT1	Protein kinase, membrane associated tyrosine/threonine 1, transcript variant 2	-4.53
	GDF15	Growth differentiation factor 15	-4.43
	LOC643031	Similar to NADH dehydrogenase subunit 5	-4.06
	MAZ	MYC-associated zinc finger protein, transcript variant 2	-3.93
	CCT7	Chaperonin containing TCP1, subunit 7 (eta), transcript variant 1	-3.91
	NGRN	Neugrin, neurite outgrowth associated (NGRN), transcript variant 1	-3.87
	LOC727761	Similar to Deoxythymidylate kinase, transcript variant 4	-3.86
	CNBP	CCHC-type zinc finger, nucleic acid binding protein	-3.85
	MORF4L2	Mortality factor 4 like 2	-3.80

doi:10.1371/journal.pone.0049045.t003

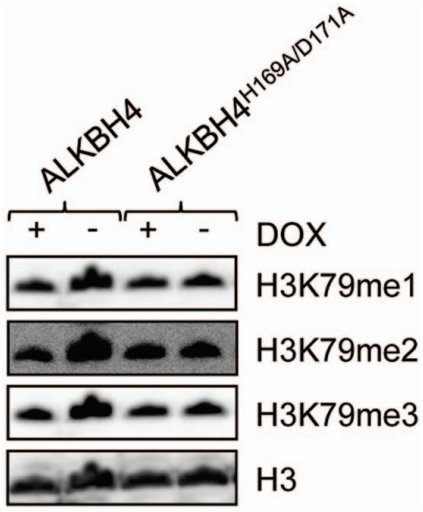


Figure 4. Ectopic ALKBH4 expression does not change H3K79 methylation levels *in vivo*. As analyzed by Western blotting, all three methylation states (mono- di- and tri-methylation) of the H3K79 residue remained similar in histones purified from stable HEK293 transfectants after doxycycline-dependent over-expression of either ALKBH4 or an enzymatically inactive mutant (ALKBH4$^{H169A/D171A}$), compared to the equivalent, non-induced cells. Signal intensities of bands corresponding to methylated histones were quantified using ImageJ [69], and normalized to the total histone H3 load, but no effect of ALKBH4 overexpression was detected (not shown).
doi:10.1371/journal.pone.0049045.g004

damage as well as in homologous recombination (HR) during meiosis, and maybe also mitosis. Thus, we speculate whether ALKBH7 might function in general regulation of cellular responses to repair of DNA double strand breaks by homologous recombination.

Conclusions

In the current work we have identified several partners of the ALKBH4 protein and an analysis of these partners suggests some processes, such as chromatin regulation and transcription, in which ALKBH4 may play an important part. Moreover, through analysis of global gene expression changes in response to ALKBH7 over-expression, we observed enrichment of differentially expressed genes in processes like cell cycle and spermatogenesis, supporting the previous annotation of ALKBH7 as a cell proliferation and spermatogenesis-related protein. Obviously, more rigorous studies are required to firmly establish the possible role of ALKBH4 and ALKBH7 in these processes, but the present study may represent a useful starting point.

Materials and Methods

Yeast Two-hybrid Assays

Two libraries (Human Placenta RP4 and Human Fetal Brain RP1) were screened using full-length human ALKBH4 as bait (N-LexA-ALKBH4-C fusion). Additionally, an enzymatically inactive ALKBH4 mutant (ALKBH4$^{H169A/D171A}$; mutagenesis described in [30]) was used as bait to screen the human placenta RP4 library. All screens and subsequent data analysis were performed by Hybrigenics.

Plasmid Construction

For generation of ALKBH4-ECFP and -EYFP fusion constructs, human ALKBH4 cDNA was subcloned between the EcoRI and BamHI sites of the pECFP-N1 and pEYFP-N1 vectors (Clontech), respectively. To prepare the AF9-EYFP plasmid, *AF9* (IMAGE-5298142) was PCR amplified (primers *AF9-fwd* and *AF9-rev*, primer sequences can be found in Table S1) and cloned into the XhoI/KpnI site of pEYFP-N1. pSXG-p300BP [41] was used for amplification of the bromodomain and PHD finger encoding fragment of p300 (primers $p300_{BP}$*-fwd* and $p300_{BP}$*-rev*), which was cloned into the XhoI/EcoRI site of pEYFP-C1 (Clontech), generating the EYFP-$p300_{BP}$ plasmid. To ensure nuclear access of the EYFP-$p300_{BP}$ fusion protein, a sequence encoding a nuclear localization signal (NLS) was subsequently PCR amplified from pCMV-nucEGFP-BP-HA (kindly provided by R. Aasland) using primers *NLS-fwd* and *NLS-rev*, and cloned into the EcoRI/KpnI site of the EYFP-$p300_{BP}$ plasmid. The ENL open reading frame was amplified from pBSIISK+ENL (kindly provided by R. Slany), using primers *ENL-fwd* and *ENL-rev*, and cloned into the XhoI/KpnI site of pEYFP-N1 (Clontech), generating the ENL-EYFP construct. A fragment encoding the N-terminus of ENL (aa 1–141) was amplified (primers *ENL-fwd* and *YEATS-rev*) and cloned into the XhoI/KpnI site of pEYFP-C1, resulting in the EYFP-ENL_{YEATS} construct, encoding a YEATS domain fusion. The XhoI/KpnI site was also used for cloning of the ENL_C-EYFP fusion construct (primers ENL_C*-fwd* and *ENL-rev*), in which the YEATS domain encoding part was deleted (aa 142–559).

For generation of inducible, stable cell lines expressing either ALKBH4, ALKBH4$^{H169A/D171A}$ or ALKBH7 (described below), the respective cDNAs were amplified (primer sequences can be found in Table S1) and cloned between the BamHI and EcoRV sites of pcDNA5/FRT/TO (Invitrogen), resulting in the plasmids pcDNA5/FRT/TO-ALKBH4, pcDNA5/FRT/TO-ALKBH4$^{H169A/D171A}$ and pcDNA5/FRT/TO-ALKBH7. A FLAG epitope tag was introduced in the reverse primers for simple detection of the proteins. All constructs were verified by DNA sequencing.

In vivo Acetylation Assay

The assay was performed essentially as previously described in [68]. Briefly, plasmids encoding FLAG-ALKBH4 (pCIneoB-3xFLAG-ALKBH4) and p300 (pCMVβ-p300) were transiently co-transfected into HCT116 cells using FuGene6 (Roche) according to the manufacturer's specifications. Empty vectors were used as controls. Twenty four hours after transfection, the cells were treated with the deacetylase inhibitor Trichostatin A (TSA, 2 μM) for 30 minutes or left untreated. Total cell extracts were prepared in lysis buffer (50 mM HEPES pH 7.9, 420 mM NaCl, 0.5% NP-40, 1 mM phenylmethylsulfonyl fluoride (PMSF), 0.5 mM dithiothreitol (DTT) and Complete protease inhibitor cocktail (Roche)). Subsequently, 1–2 mg total cell extract and 2–5 μg anti-FLAG antibody (Sigma F3165) was used for immunoprecipitation of FLAG-ALKBH4. The resulting precipitates were analyzed by Western blotting using an anti-acetyl-lysine antibody (Upstate 06-933) followed by reprobing of the membrane with the anti-FLAG antibody.

GST Pull-down Assay

The TNT T7 Quick for PCR DNA system (Promega) was used according to the manufacturer's descriptions to produce *in vitro* transcribed/translated [^{35}S]-methionine labelled ALKBH4 in

presence of Redivue L-[^{35}S]-methionine (GE Healthcare). To minimize production of truncated protein products due to partly degraded mRNA, 10–15 μg RibonucleaseA (Sigma) was added to the reaction mixtures subsequent to translation and further incubated at room temperature for 10 minutes. GST fusions of ENL and p300$_{BP}$ were separately expressed in *E. coli*, immobilized on Glutathione Sepharose resin (GE Healthcare) and washed with interaction buffer (50 mM Tris-HCl pH 8.5, 12 mM NaCl, 0.1 mM ZnAc, 150 mM KCl, 2 mM $MgCl_2$, 10 mM 2-mercaptoethanol, 0.1% Triton X-100). *In vitro* translation reaction mixture (5–10 μl) was added to the immobilized GST-fusion proteins, which were incubated with gentle agitation for 30 min at room temperature. Reactions were subsequently subjected to four washes with interaction buffer and one wash with 50 mM Tris-HCl pH 8.0 prior to elution with GST-elution buffer (50 mM Tris-HCl pH 8.0, 15 mM reduced glutathione, 0.1 mM ZnAc). Similarly, reciprocal pull-downs using GST-ALKBH4 and [^{35}S]-labelled p300$_{BP}$ (alternatively full-length p300) were also performed. GST-only was included as control. Eluted proteins were analyzed by SDS-PAGE followed by phosphor imaging using LE Storage Phosphor Screens (GE Healthcare) and an exposure time of 1–4 days. Signals were visualized in a Typhoon^{TM} 9400 scanner (GE Healthcare).

Co-immunoprecipitation

Plasmids encoding HA-tagged versions of ENL, AF9 or p300$_{BP}$, or the empty vector (pCMV-script, Stratagene), were separately transfected into Flp-In-293-ALKBH4 cells (described below) using FuGene6 (Roche) according to the manufacturer's descriptions. Cells were treated with 2 μg/ml DOX to induce expression of FLAG-tagged ALKBH or left untreated. After 18–24 hours, the cells were harvested in phosphate-buffered saline (PBS) and washed once in PBS before $5–6\times10^6$ cells were subjected to crosslinking in PBS with 0.125–0.25% formaldehyde for 20 minutes at 37°C. The crosslinking reaction was stopped by the addition of 0.15 M glycine followed by incubation on ice for 3 minutes and subsequently at room temperature for 2 minutes.

To prepare cell extracts, the cells were washed twice with PBS and resuspended in 3 volumes (~200 μl) of Buffer I (20 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 20% glycerol, 0.5% NP-40, 1 mM DTT, Complete protease inhibitor cocktail (Roche)) containing 2 μl Omnicleave Endonuclease (200 U/μl Epicentre Technologies) followed by sonication, addition of DNase/RNase cocktail (200 U/μl Omnicleave Endonuclease, 250 U/ml Benzonase (Novagen), 30 mg/ml RNase (Sigma-Aldrich), 10 U/μl DNase (Roche), 100–300 U/mg micrococcal nuclease (Sigma-Aldrich)), incubation with gentle agitation for 1 hour at room temperature and subsequently over night at 4°C. The extracts were cleared by centrifugation at 14,000×*g* at 4°C for 10 minutes.

Immunoprecipitation was performed by pre-incubating the cell extracts with 2–5 μg anti-FLAG antibody (Sigma F3165) in 5 ml Buffer II (20 mM HEPES pH 7.9, 1.5 mM $MgCl_2$, 100 mM KCl, 0.2 mM EDTA, 10% glycerol, 1 mM DTT, Complete protease inhibitor cocktail (Roche)) at room temperature for 1 hour. Subsequently, 50 μl 10% Sepharose rec-Protein G (Invitrogen) was added and the samples were incubated with gentle agitation at 4°C for 1 hour. After extensive washing with Buffer II the immunocomplexes were resuspended in Laemmli buffer. Crosslinks were reversed by incubating the samples at 70°C for 10 minutes followed by addition of 0.1 M DTT and further incubation at 95°C for 30 minutes. Immunocomplexes were analyzed by SDS-PAGE and Western blotting using an anti-HA antibody (Abcam Ab9110), and membranes were reprobed with the anti-FLAG antibody.

Confocal Imaging

Live HeLa S3 cells (ATCC) were examined 16–24 h after transient transfection (using FuGene 6 or FuGene HD (Roche) according to the manufacturer's recommendations) of the ECFP/EYFP/RFP fusion constructs. The fluorescent images were acquired using a Zeiss LSM 510 Meta laser scanning microscope equipped with a Plan-Apochromate 63×/1.4 oil immersion objective. The images were acquired in the growth medium of the cell, with the stage heated to 37°C, using the Zeiss LSM 510 software. ECFP was excited at $\lambda = 458$ nm and detected at $\lambda = 470–500$ nm and EYFP was excited at $\lambda = 514$ nm and detected at $\lambda = 530–600$ nm. The thickness of the slice was 1 μm. All images were acquired with consecutive scans to avoid bleed through. No image processing, except contrast and intensity adjustments, were performed.

Establishment and Maintenance of Stable, Inducible Cell Lines

Inducible cell lines for stable over-expression of FLAG epitope-tagged versions of ALKBH4, ALKBH4$^{H169/D171A}$ and ALKBH7, respectively, were generated using the Flp-In^{TM} T-REx^{TM} System (Invitrogen) and the Flp-In^{TM} T-REx^{TM}-293 host cell line (Invitrogen, R780-07), according to the manufacturer's specifications. This system ensures isogenic cDNA expression from a single transcriptionally active genomic locus. Briefly, pcDNA5/FRT/TO-ALKBH4, pcDNA5/FRT/TO- ALKBH4$^{H169/D171A}$ and pcDNA5/FRT/TO-ALKBH7 (described above) were independently co-transfected with the Flp recombinase-encoding pOGG44 vector into the Flp-In^{TM} T-REx^{TM}-293 host cell line, using FuGene6 transfection agent (Roche). Selection of transfectants was performed in presence of 200 μg/ml Hygromycin B (Clontech). Single colonies conferring hygromycin B resistance and zeocin sensitivity were expanded and treated with doxycycline (Clontech). Total cell extracts were prepared using RIPA lysis buffer (Santa Cruz), according to standard methods, and screened for ALKBH over-expression by Western blotting (described below). Stable transfectants were maintained in Dulbecco's modified Eagle's medium (DMEM) (Lonza) supplemented with 10% tetracycline-free fetal bovine serum (FBS) (Clontech), 100 U/ml penicillin (Lonza), 100 U/ml streptomycin (Lonza), 2 mM L-glutamine (Lonza), 15 μg/ml blasticidin-S (Invitrogen) and 200 μg/ml hygromycin B. Ectopic ALKBH expression was induced by addition of 2 μg/ml doxycycline (DOX) to the medium. For microarray analysis, transgene expression was induced at a cell confluence level of approximately 50%. Cells were harvested 48 hours after induction. All samples were prepared in triplicates.

Histone Purification

Histones were purified from Flp-In^{TM}-293 cells (Invitrogen) containing a stable, doxycycline-inducible integration expressing FLAG-ALKBH4 or FLAG-ALKBH4$^{H169A/D171A}$ (described above), either treated with doxycycline or untreated, using the Histone purification mini kit (Active Motif) according to the manufacturer's specifications. Briefly, cells were lysed in Extraction Buffer at 4°C for 1 hour. Cleared lysates were neutralized and loaded onto pre-equilibrated spin columns which were washed prior to histone elution. Purified histones were subsequently concentrated by perchloric acid precipitation.

Antibodies and Western Blot Analysis

Total cell extracts or histones were subjected to separation by SDS-PAGE (NuPAGE® SDS-PAGE Gel System, Invitrogen) and transferred onto Invitrolon PVDF membranes (Invitrogen). Membranes were subjected to blocking with 5% dry milk in PBS with 0.1% Tween-20 (PBS-T) for 1 hour. After incubation with primary antibody diluted in PBS-T with 5% dry milk for 1 hour, the membranes were washed three times with PBS-T, incubated with secondary antibody diluted in PBS-T with 2,5% dry milk for 1 hour and subjected to three additional washes with PBS-T. Primary antibodies used were anti-FLAG (Sigma F3165), anti-HA (Abcam Ab9110), anti-GAPDH (Applied Biosystems AM4300), anti-acetyl-lysine (Upstate 06-933) anti-H3K79me1 (Abcam ab2886), anti-H3K79me2 (Abcam ab3594), anti-H3K79me3 (Abcam ab2621) and anti-H3 (Abcam ab1791). While the anti-H3K79me1 and anti-H3K79me3 antibodies were specific towards mono- and trimethylated H3K79, respectively, anti-H3K79me2 antibody specificity towards dimethylated H3K79 was ensured by pre-incubating the antibody with two H3 derived peptides containing either mono- (Abcam ab4555) or trimethylated (Abcam ab4557) H3K79 (1 μg/ml of each) for 30 min at RT before incubation with the membrane, thereby blocking mono- and trimethyl reactive sites (data not shown). Anti-FLAG and anti-GAPDH antibodies were used in combination with an alkaline phosphatase-conjugated secondary antibody and the Amersham ECF detection system (GE Healthcare). Fluorescence signal detection was performed in a Typhoon scanner 9400 (GE Healthcare). The remaining antibodies were used in combination with a horseradish peroxidase-conjugated secondary antibody and the SuperSignal West Dura kit (Thermo Scientific). Chemiluminescence signals were visualized on a Kodak Image Station 4000R Pro instrument (Carestream Health). Densitometry was performed using the ImageJ software [69].

RNA and Genomic DNA Isolation

Total RNA and genomic DNA were isolated from cell lines over-expressing ALKBH4 or ALKBH7 or the equivalent non-induced cell lines. For RNA isolation, TRIzol (Invitrogen) was used according to the manufacturer's descriptions. The RNA was subsequently subjected to a clean-up step using the RNeasy Mini Kit (Qiagen) according to the RNA Cleanup procedure in the supplied manual. Genomic DNA was isolated using the QIAamp DNA Mini Kit (Qiagen) according to the manufacturer's protocol.

Quantitative PCR Analysis

SuperScript II Reverse Transcriptase (Invitrogen) was used to synthesize cDNA, using oligo$(dT)_{12-18}$ primers (Invitrogen) and 5 μg of total RNA. cDNA synthesis was performed according to the manufacturer's specifications, except that the incubation step prior to addition of the SuperScript II RT was omitted. Quantitative PCR (qPCR) was performed in 20 μl reactions using SYBR Green master mix (Qiagen), 4 μl cDNA (diluted 1:16), and 10 pmol *ALKBH4*-specific primers (*ALKBH4-qPCR-fwd* and *ALKBH4-qPCR-rev*, primer sequences can be found in Table S1). Reactions were performed on a LightCycler 1.5 instrument (Roche). LightCycler software 3.5.3 (Roche) was used for data analysis and quantification was performed using the ΔC_T method [70] with *β-actin* as endogenous reference (primers *beta-actin-fwd* and *beta-actin-rev*). All experiments were performed in triplicates.

Microarrays and Data Analysis

mRNA expression and CpG methylation profiling was performed by the Helse Sør-Øst/University of Oslo Genomics Core Facility using the Illumina HumanWG-6 v3 Expression BeadChip and Illumina Infinium HumanMethylation27 BeadChip, respectively, according to the manufacturer's protocols. Data extraction and initial quality control of the bead summary raw data were performed using GenomeStudio v2011.1 from Illumina and the Gene Expression and Methylation module v1.9.0. The data was annotated using the HumanWG-6_V3_0_R3_11282955_A and HumanMethylation27_270596_v1.2 annotation files from Illumina. Unnormalized bead intensities were exported into a tab delimited text file and imported into J-express Pro (v. 2.7) for further downstream analysis [71]. Bead intensities were quantile normalized, and rank product analysis was performed to identify gene expression changes [72]. A significant threshold of q-value <5 and a fold-change larger than 1.35 for ALKBH4 or 2 for ALKBH7 was used to identify differentially expressed genes. Differentially expressed genes were further analyzed in MetaCore (GeneGo) to identify functional enrichment.

CpG methylation data was analyzed using the GenomeStudio Methylation Module, where avgBeta (average ratio of signal from methylated probe relative to the sum of both methylated and unmethylated probes) was calculated and used for comparisons within sample groups.

The mRNA expression and CpG methylation datasets have been deposited in the GEO data repository (www.ncbi.nlm.nih.gov/geo/, accession number GSE39135).

Supporting Information

Figure S1 Partial co-localization of ALKBH4 and ENL with RNA Polymerase I subunit RPA43 in nucleolar speckles. Co-expression of ALKBH4-EYFP with ECFP-ENL and RPA43-RFP in HeLa cells, as analyzed by confocal fluorescence microscopy. Insets are enlargements of boxed areas.
(PDF)

Figure S2 Effects of ALKBH4 and ALKBH7 over-expression on the global DNA methylation pattern. CpG methylation profiles were analyzed in stably transfected HEK293 cells before vs. after doxycycline induced over-expression of ALKBH4 (left panel) or ALKBH7 (right panel), using the Illumina Infinium HumanMethylation27 BeadChip. DOX, doxycycline.
(PDF)

Table S1 Primer sequences.
(PDF)

Acknowledgments

We are grateful to R. Aasland and R. Slany for providing the pCMV-nucEGFP-BP-HA and pBSIISK+ENL plasmids, respectively. We also thank M. Kjos for technical assistance.

Author Contributions

Conceived and designed the experiments: LGB TJM MO LAMZ PØF. Performed the experiments: LGB TJM MO SMO. Analyzed the data: LGB TJM MO SMO LAMZ PØF. Wrote the paper: LGB PØF.

References

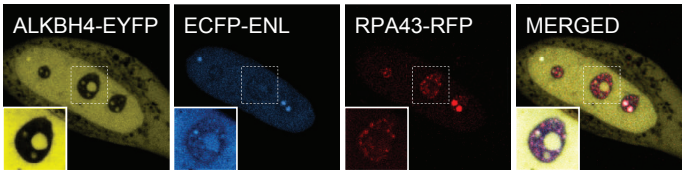
1. Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, et al. (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. Science 324: 930–935.

2. Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, et al. (2006) Histone demethylation by a family of JmjC domain-containing proteins. Nature 439: 811–816.

3. Roach PL, Clifton IJ, Fulop V, Harlos K, Barton GJ, et al. (1995) Crystal structure of isopenicillin N synthase is the first from a new structural family of enzymes. Nature 375: 700–704.
4. Myllyharju J (2003) Prolyl 4-hydroxylases, the key enzymes of collagen biosynthesis. Matrix Biol 22: 15–24.
5. Epstein AC, Gleadle JM, McNeill LA, Hewitson KS, O'Rourke J, et al. (2001) C. elegans EGL-9 and mammalian homologs define a family of dioxygenases that regulate HIF by prolyl hydroxylation. Cell 107: 43–54.
6. Lando D, Peet DJ, Gorman JJ, Whelan DA, Whitelaw ML, et al. (2002) FIH-1 is an asparaginyl hydroxylase enzyme that regulates the transcriptional activity of hypoxia-inducible factor. Genes Dev 16: 1466–1471.
7. Aravind L, Koonin EV (2001) The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol 2: RESEARCH0007.
8. Falnes PO, Johansen RF, Seeberg E (2002) AlkB-mediated oxidative demethylation reverses DNA damage in Escherichia coli. Nature 419: 178–182.
9. Trewick SC, Henshaw TF, Hausinger RP, Lindahl T, Sedgwick B (2002) Oxidative demethylation by Escherichia coli AlkB directly reverts DNA base damage. Nature 419: 174–178.
10. Koivisto P, Duncan T, Lindahl T, Sedgwick B (2003) Minimal methylated substrate and extended substrate range of Escherichia coli AlkB protein, a 1-methyladenine-DNA dioxygenase. J Biol Chem 278: 44348–44354.
11. Delaney JC, Smeester L, Wong C, Frick LE, Taghizadeh K, et al. (2005) AlkB reverses etheno DNA lesions caused by lipid oxidation in vitro and in vivo. Nat Struct Mol Biol 12: 855–860.
12. Mishina Y, Yang CG, He C (2005) Direct repair of the exocyclic DNA adduct 1,N6-ethenoadenine by the DNA repair AlkB proteins. J Am Chem Soc 127: 14594–14595.
13. Frick LE, Delaney JC, Wong C, Drennan CL, Essigmann JM (2007) Alleviation of 1,N6-ethanoadenine genotoxicity by the Escherichia coli adaptive response protein AlkB. Proc Natl Acad Sci U S A 104: 755–760.
14. Aas PA, Otterlei M, Falnes PO, Vagbo CB, Skorpen F, et al. (2003) Human and bacterial oxidative demethylases repair alkylation damage in both RNA and DNA. Nature 421: 859–863.
15. Kurowski MA, Bhagwat AS, Papaj G, Bujnicki JM (2003) Phylogenomic identification of five new human homologs of the DNA repair enzyme AlkB. BMC Genomics 4: 48.
16. Gerken T, Girard CA, Tung YC, Webby CJ, Saudek V, et al. (2007) The obesity-associated FTO gene encodes a 2-oxoglutarate-dependent nucleic acid demethylase. Science 318: 1469–1472.
17. Jia G, Yang CG, Yang S, Jian X, Yi C, et al. (2008) Oxidative demethylation of 3-methylthymine and 3-methyluracil in single-stranded DNA and RNA by mouse and human FTO. FEBS Lett 582: 3313–3319.
18. Jia G, Fu Y, Zhao X, Dai Q, Zheng G, et al. (2011) N6-methyladenosine in nuclear RNA is a major substrate of the obesity-associated FTO. Nat Chem Biol 7: 885–887.
19. Songe-Moller L, van den Born E, Leihne V, Vagbo CB, Kristoffersen T, et al. (2010) Mammalian ALKBH8 possesses tRNA methyltransferase activity required for the biogenesis of multiple wobble uridine modifications implicated in translational decoding. Mol Cell Biol 30: 1814–1827.
20. van den Born E, Vagbo CB, Songe-Moller L, Leihne V, Lien GF, et al. (2011) ALKBH8-mediated formation of a novel diastereomeric pair of wobble nucleosides in mammalian tRNA. Nat Commun 2: 172.
21. Fu D, Brophy JA, Chan CT, Atmore KA, Begley U, et al. (2010) Human AlkB homolog ABH8 Is a tRNA methyltransferase required for wobble uridine modification and DNA damage survival. Mol Cell Biol 30: 2449–2459.
22. Fu Y, Dai Q, Zhang W, Ren J, Pan T, et al. (2010) The AlkB domain of mammalian ABH8 catalyzes hydroxylation of 5-methoxycarbonylmethyluridine at the wobble position of tRNA. Angew Chem Int Ed Engl 49: 8885–8888.
23. Westbye MP, Feyzi E, Aas PA, Vagbo CB, Talstad VA, et al. (2008) Human AlkB homolog 1 is a mitochondrial protein that demethylates 3-methylcytosine in DNA and RNA. J Biol Chem 283: 25046–25056.
24. Pan Z, Sikandar S, Witherspoon M, Dizon D, Nguyen T, et al. (2008) Impaired placental trophoblast lineage differentiation in Alkbh1(−/−) mice. Dev Dyn 237: 316–327.
25. Nordstrand LM, Svard J, Larsen E, Nilsen A, Ougland R, et al. (2010) Mice lacking Alkbh1 display sex-ratio distortion and unilateral eye defects. PLoS One 5: e13827.
26. Sedgwick B, Bates PA, Paik J, Jacobs SC, Lindahl T (2007) Repair of alkylated DNA: recent advances. DNA Repair (Amst) 6: 429–442.
27. Loenarz C, Schofield CJ (2008) Expanding chemical biology of 2-oxoglutarate oxygenases. Nat Chem Biol 4: 152–156.
28. Korvald H, Molstad Moe AM, Cederkvist FH, Thiede B, Laerdahl JK, et al. (2011) Schizosaccharomyces pombe Ofd2 is a nuclear 2-oxoglutarate and iron dependent dioxygenase interacting with histones. PLoS One 6: e25188.
29. Thalhammer A, Bencokova Z, Poole R, Loenarz C, Adam J, et al. (2011) Human AlkB homologue 5 is a nuclear 2-oxoglutarate dependent oxygenase and a direct target of hypoxia-inducible factor 1alpha (HIF-1alpha). PLoS One 6: e16210.
30. Bjornstad LG, Zoppellaro G, Tomter AB, Falnes PO, Andersson KK (2011) Spectroscopic and magnetic studies of wild-type and mutant forms of the Fe(II)- and 2-oxoglutarate-dependent decarboxylase ALKBH4. Biochem J 434: 391–398.
31. Ogryzko VV, Schiltz RL, Russanova V, Howard BH, Nakatani Y (1996) The transcriptional coactivators p300 and CBP are histone acetyltransferases. Cell 87: 953–959.
32. Miura Y, Tam T, Ido A, Morinaga T, Miki T, et al. (1995) Cloning and characterization of an ATBF1 isoform that expresses in a neuronal differentiation-dependent manner. J Biol Chem 270: 26840–26848.
33. Morinaga T, Yasuda H, Hashimoto T, Higashio K, Tamaoki T (1991) A human alpha-fetoprotein enhancer-binding protein, ATBF1, contains four homeodomains and seventeen zinc fingers. Mol Cell Biol 11: 6041–6049.
34. Nakai A, Tanabe M, Kawazoe Y, Inazawa J, Morimoto RI, et al. (1997) HSF4, a new member of the human heat shock factor family which lacks properties of a transcriptional activator. Mol Cell Biol 17: 469–481.
35. Tanabe M, Sasai N, Nagata K, Liu XD, Liu PC, et al. (1999) The mammalian HSF4 gene generates both an activator and a repressor of heat shock genes by alternative splicing. J Biol Chem 274: 27845–27856.
36. Nakamura T, Alder H, Gu Y, Prasad R, Canaani O, et al. (1993) Genes on chromosomes 4, 9, and 19 involved in 11q23 abnormalities in acute leukemia share sequence homology and/or common motifs. Proc Natl Acad Sci U S A 90: 4631–4635.
37. Rubnitz JE, Morrissey J, Savage PA, Cleary ML (1994) ENL, the gene fused with HRX in t(11;19) leukemias, encodes a nuclear protein with transcriptional activation potential in lymphoid and myeloid cells. Blood 84: 1747–1752.
38. Mueller D, Bach C, Zeisig D, Garcia-Cuellar MP, Monroe S, et al. (2007) A role for the MLL fusion partner ENL in transcriptional elongation and chromatin modification. Blood 110: 4445–4454.
39. Srinivasan RS, de Erkenez AC, Hemenway CS (2003) The mixed lineage leukemia fusion partner AF9 binds specific isoforms of the BCL-6 corepressor. Oncogene 22: 3395–3406.
40. Tsujikawa K, Koike K, Kitae K, Shinkawa A, Arima H, et al. (2007) Expression and sub-cellular localization of human ABH family molecules. J Cell Mol Med 11: 1105–1116.
41. Ragvin A, Valvatne H, Erdal S, Arskog V, Tufteland KR, et al. (2004) Nucleosome binding by the bromodomain and PHD finger of the transcriptional cofactor p300. J Mol Biol 337: 773–788.
42. Hirschler-Laszkiewicz I, Cavanaugh A, Hu Q, Catania J, Avantaggiati ML, et al. (2001) The role of acetylation in rDNA transcription. Nucleic Acids Res 29: 4114–4124.
43. Bishop DK, Park D, Xu L, Kleckner N (1992) DMC1: a meiosis-specific yeast homolog of E. coli recA required for recombination, synaptonemal complex formation, and cell cycle progression. Cell 69: 439–456.
44. Dray E, Siaud N, Dubois E, Doutriaux MP (2006) Interaction between Arabidopsis Brca2 and its partners Rad51, Dmc1, and Dss1. Plant Physiol 140: 1059–1069.
45. Sharma GG, So S, Gupta A, Kumar R, Cayrou C, et al. (2010) MOF and histone H4 acetylation at lysine 16 are critical for DNA damage response and double-strand break repair. Mol Cell Biol 30: 3582–3595.
46. Ooga M, Inoue A, Kageyama S, Akiyama T, Nagata M, et al. (2008) Changes in H3K79 methylation during preimplantation development in mice. Biol Reprod 78: 413–424.
47. Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, et al. (2002) Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. Curr Biol 12: 1052–1058.
48. Xu W, Yang H, Liu Y, Yang Y, Wang P, et al. (2011) Oncometabolite 2-hydroxyglutarate is a competitive inhibitor of alpha-ketoglutarate-dependent dioxygenases. Cancer Cell 19: 17–30.
49. Dhalluin C, Carlson JE, Zeng L, He C, Aggarwal AK, et al. (1999) Structure and ligand of a histone acetyltransferase bromodomain. Nature 399: 491–496.
50. Musselman CA, Kutateladze TG (2011) Handpicking epigenetic marks with PHD fingers. Nucleic Acids Res 39: 9061–9071.
51. Aasland R, Gibson TJ, Stewart AF (1995) The PHD finger: implications for chromatin-mediated transcriptional regulation. Trends Biochem Sci 20: 56–59.
52. Zeng L, Yap KL, Ivanov AV, Wang X, Mujtaba S, et al. (2008) Structural insights into human KAP1 PHD finger-bromodomain and its role in gene silencing. Nat Struct Mol Biol 15: 626–633.
53. Agricola E, Randall RA, Gaarenstroom T, Dupont S, Hill CS (2011) Recruitment of TIF1gamma to chromatin via its PHD finger-bromodomain activates its ubiquitin ligase and transcriptional repressor activities. Mol Cell 43: 85–96.
54. Bordoli L, Husser S, Luthi U, Netsch M, Osmani H, et al. (2001) Functional analysis of the p300 acetyltransferase domain: the PHD finger of p300 but not of CBP is dispensable for enzymatic activity. Nucleic Acids Res 29: 4462–4471.
55. Zeisig DT, Bittner CB, Zeisig BB, Garcia-Cuellar MP, Hess JL, et al. (2005) The eleven-nineteen-leukemia protein ENL connects nuclear MLL fusion partners with chromatin. Oncogene 24: 5525–5532.
56. He N, Chan CK, Sobhian B, Chou S, Xue Y, et al. (2011) Human Polymerase-Associated Factor complex (PAFc) connects the Super Elongation Complex (SEC) to RNA polymerase II on chromatin. Proc Natl Acad Sci U S A 108: E636–645.
57. Schulze JM, Wang AY, Kobor MS (2010) Reading chromatin: insights from yeast into YEATS domain structure and function. Epigenetics 5: 573–577.
58. Wang AY, Schulze JM, Skordalakes E, Gin JW, Berger JM, et al. (2009) Asf1-like structure of the conserved Yaf9 YEATS domain and role in H2A.Z deposition and acetylation. Proc Natl Acad Sci U S A 106: 21573–21578.

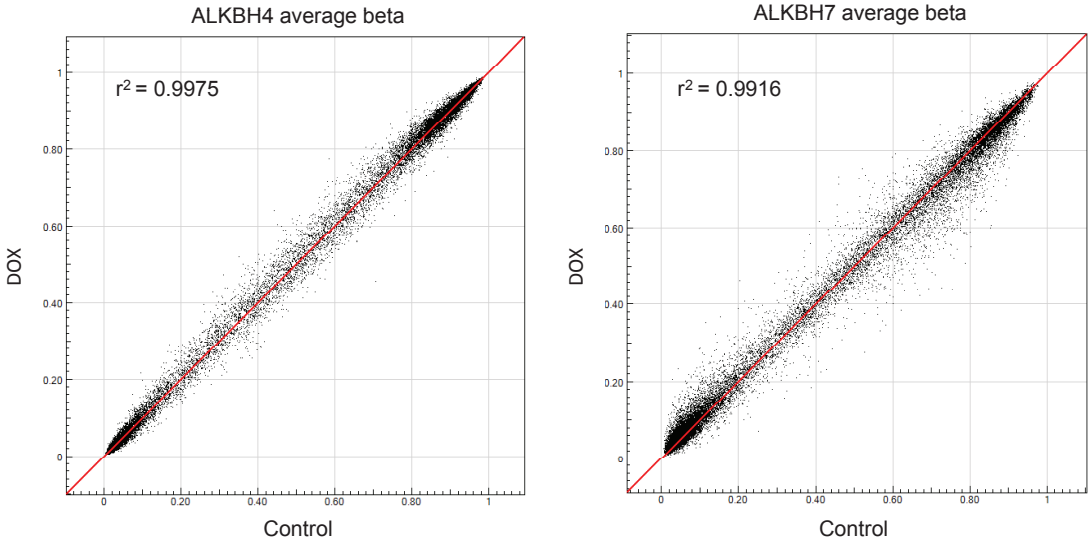
59. Brayer KJ, Segal DJ (2008) Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell Biochem Biophys 50: 111–131.
60. Nojiri S, Joh T, Miura Y, Sakata N, Nomura T, et al. (2004) ATBF1 enhances the suppression of STAT3 signaling by interaction with PIAS3. Biochem Biophys Res Commun 314: 97–103.
61. Ida K, Kitabayashi I, Taki T, Taniwaki M, Noro K, et al. (1997) Adenoviral E1A-associated protein p300 is involved in acute myeloid leukemia with t(11;22)(q23;q13). Blood 90: 4699–4704.
62. Iida S, Seto M, Yamamoto K, Komatsu H, Tojo A, et al. (1993) MLLT3 gene on 9p22 involved in t(9;11) leukemia encodes a serine/proline rich protein homologous to MLLT1 on 19p13. Oncogene 8: 3085–3092.
63. Krivtsov AV, Feng Z, Lemieux ME, Faber J, Vempati S, et al. (2008) H3K79 methylation profiles define murine and human MLL-AF4 leukemias. Cancer Cell 14: 355–368.
64. Slany RK (2009) The molecular biology of mixed lineage leukemia. Haematologica 94: 984–993.
65. Zhang K, Sridhar VV, Zhu J, Kapoor A, Zhu JK (2007) Distinctive core histone post-translational modification patterns in Arabidopsis thaliana. PLoS One 2: e1210.
66. Kivirikko KI, Myllyla R, Pihlajaniemi T (1989) Protein hydroxylation: prolyl 4-hydroxylase, an enzyme with four cosubstrates and a multifunctional subunit. FASEB J 3: 1609–1617.
67. Lee DH, Jin SG, Cai S, Chen Y, Pfeifer GP, et al. (2005) Repair of methylation damage in DNA and RNA by mammalian AlkB homologues. J Biol Chem 280: 39448–39459.
68. Buerki C, Rothgiesser KM, Valovka T, Owen HR, Rehrauer H, et al. (2008) Functional relevance of novel p300-mediated lysine 314 and 315 acetylation of RelA/p65. Nucleic Acids Res 36: 1665–1680.
69. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. Nat Methods 9: 671–675.
70. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402–408.
71. Dysvik B, Jonassen I (2001) J-Express: exploring gene expression data using Java. Bioinformatics 17: 369–370.
72. Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 573: 83–92.
73. Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, et al. (2012) InterPro in 2011: new developments in the family and domain prediction database. Nucleic Acids Res 40: D306–312.
74. Sigrist CJ, Cerutti L, de Castro E, Langendijk-Genevaux PS, Bulliard V, et al. (2010) PROSITE, a protein domain database for functional characterization and annotation. Nucleic Acids Res 38: D161–166.
75. Formstecher E, Aresta S, Collura V, Hamburger A, Meil A, et al. (2005) Protein interaction mapping: a Drosophila case study. Genome Res 15: 376–384.

Bjørnstad *et al.* – Supplementary Figure S1



Supplementary Figure S1. Partial co-localization of ALKBH4 and ENL with RNA Polymerase I subunit RPA43 in nucleolar speckles. Co-expression of ALKBH4-EYFP with ECFP-ENL and RPA43-RFP in HeLa cells, as analyzed by confocal fluorescence microscopy. Insets are enlargements of boxed areas.

Bjørnstad *et al.* – Supplementary Figure S2



Supplementary Figure S2. Effects of ALKBH4 and ALKBH7 over-expression on the global DNA methylation pattern. CpG methylation profiles were analyzed in stably transfected HEK293 cells before vs. after doxycycline-induced over-expression of ALKBH4 (left panel) or ALKBH7 (right panel), using the Illumina Infinium HumanMethylation27 BeadChip. DOX, doxycycline.

Bjørnstad *et al.*- Supplementary Table S1

Supplementary Table S1. Primer sequences

Primer name	Primer sequence (5'-3')	Application
$p300_{BP}$-fwd	ATCGAGTACTCGAGTCATGGCCCAGGTGCAGCTGCA	Plasmid cloning
$p300_{BP}$-rev	ACGCTACTGAATTCCTTCTTTCCTAGTTCGTGCACTTTTCTTTA	Plasmid cloning
NLS-fwd	GTCTAGACGAATTCACGATCCAAAAAAGAAGAGAAAGG	Plasmid cloning
NLS-rev	TCAGCTAGGGTACCAGTACCTTTCTCTTCTTTTTTGG	Plasmid cloning
AF9-fwd	AGATTCGACTCGAGTAATGGCTAGCTCGTGTGCCGTGCAG	Plasmid cloning
AF9-rev	TACGACTTGGTACCGAGGATGTTCCAGATGTTTCCAGGTAACTC	Plasmid cloning
ENL-fwd	TCCGACTACTCGAGTAATGGACAATCAGTGCACCGTCCAGG	Plasmid cloning
ENL-rev	TACGACTTGGTACCGATGTGGCCACGGCCTCCAGGCA	Plasmid cloning
YEATS-rev	CGACATGGTACCGACACCCCGCCGGCCCGCAG	Plasmid cloning
ENL_C-fwd	CGACTACTCGAGTAATGGTAATGCCCGAAGGAGC	Plasmid cloning
ALKBH4-fwd	ACTACTGGGGATCCATGGCGGCGGCTGCCGCCGA	Plasmid cloning
ALKBH4-FLAG-rev	ATGTAGACGATATCTCATTTATCGTCATCGTCTTTGTAGTCGAACACGGGTCTTCCCTGGAAGGA	Plasmid cloning
ALKBH7-fwd	ACTACTGGGGATCCATGGCCGGGACTGGGCTGCTG	Plasmid cloning
ALKBH7-FLAG-rev	ATGTAGACGATATCTCATTTATCGTCATCGTCTTTGTAGTCGAAGCAGGCTGGGGGCGGCT	Plasmid cloning
ALKBH4-qPCR-fwd	TTCGGGAATGCGGTTG	qPCR
ALKBH4-qPCR-rev	CCTCGATCAGCATCACTC	qPCR
Beta-actin-fwd	CGTGGGGCGCCCCAGGCACCA	qPCR
Beta-actin-rev	TTGGCCTTGGGGTTCAGGGGGG	qPCR